GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 14:46:57 ; Search time 7316.32 Seconds
(without alignments)
4710.839 Million cell updates/sec

Title: US-09-826-581-5
Perfect score: 1647
Sequence: 1 ttggtctggggctggccaca.....acaccagcctcttagtcttc 1647

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1647	100.0	1647	6	AX281582	AX281582 Sequence
	2	1590	96.5	2290	9	HSA249977	AJ249977 Homo sapi
	3	1453	88.2	2115	6	AX099802	AX099802 Sequence
	4	1453	88.2	2115	9	AF214519	AF214519 Homo sapi
	5	1447	87.9	2109	6	AX099776	AX099776 Sequence
	6	1172.8	71.2	2022	6	AX099804	AX099804 Sequence
	7	1140.4	69.2	1873	4	AF214520	AF214520 Sus scrof
	8	1140.4	69.2	1873	6	AX099800	AX099800 Sequence
	9	1134.4	68.9	1867	6	AX099774	AX099774 Sequence
	10	414.8	25.2	1328	10	RNAMPKGAM	X95578 R.norvegicu
	11	414.8	25.2	1550	10	RNU42413	U42413 Rattus norv
	12	407.2	24.7	989	6	AX281579	AX281579 Sequence
c	13	407.2	24.7	196554	2	AC073128	AC073128 Homo sapi
	14	396.8	24.1	1623	10	AF036535	AF036535 Mus muscu
	15	395.4	24.0	3132	10	BC015283	BC015283 Mus muscu
c	16	395.2	24.0	206854	9	AC009974	AC009974 Homo sapi
	17	393.4	23.9	1578	9	HSU42412	U42412 Human 5'-AM
	18	393.4	23.9	1677	9	BC000358	BC000358 Homo sapi
	19	366.6	22.3	1167	9	AB025580	AB025580 Homo sapi
	20	366.6	22.3	1435	6	AR139104	AR139104 Sequence
	21	366.6	22.3	2194	9	AF087875	AF087875 Homo sapi
	22	366.6	22.3	2203	9	BC020540	BC020540 Homo sapi
	23	366.6	22.3	2223	9	AK001887	AK001887 Homo sapi
	24	363.4	22.1	2062	9	HSA249976	AJ249976 Homo sapi
	25	300.8	18.3	2082	3	AF094764	AF094764 Drosophil
	26	299.2	18.2	4666	3	AY070541	AY070541 Drosophil
	27	278.8	16.9	1014	6	AX281581	AX281581 Sequence
	28	278.8	16.9	152129	2	AC027416	AC027416 Homo sapi
	29	278.8	16.9	196554	2	AC073128	AC073128 Homo sapi
	30	273.6	16.6	5888	4	AF214521	AF214521 Sus scrof
c	31	267.4	16.2	227724	2	AF336381	AF336381 Mus muscu
c	32	198.2	12.0	152129	2	AC027416	AC027416 Homo sapi
	33	168.2	10.2	1722	6	AX281580	AX281580 Sequence
	34	131	8.0	3497	3	AF094763	AF094763 Drosophil
	35	127.2	7.7	80069	2	AC019671	AC019671 Drosophil
	36	127.2	7.7	173634	3	AC009344	AC009344 Drosophil
	37	127.2	7.7	195868	3	AC008308	AC008308 Drosophil
	38	127.2	7.7	230266	3	AE003733	AE003733 Drosophil
	39	91.4	5.5	106	11	G67375	G67375 A898 Human
	40	84.8	5.1	151	6	AX305717	AX305717 Sequence
	41	83	5.0	142903	2	AL627254	AL627254 Danio rer
c	42	78.4	4.8	110000	2	CEY111B2_1	Continuation (2 of
c	43	78.4	4.8	110000	2	CEY111B2_2	Continuation (3 of
c	44	78.4	4.8	257702	3	CEY111B2A	AL132904 Caenorhab
	45	75	4.6	14411	4	AF329081	AF329081 Bos tauru

ALIGNMENTS

RESULT 1
AX281582
LOCUS AX281582 1647 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 5 from Patent WO0177305.
ACCESSION AX281582
VERSION AX281582.1 GI:16608833
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Andersson,L., Luthman,H. and Marklund,S.
TITLE Variants of the human amp-activated protein kinase gamma 3 subunit
JOURNAL Patent: WO 0177305-A 5 18-OCT-2001;
Arexis AB (SE)
FEATURES Location/Qualifiers
source 1. .1647
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS 20. .1489
/note="unnamed protein product"
/codon_start=1

```
                     /protein_id="CAD10589.1"
                     /db_xref="GI:16608834"
                     /translation="MEPGLEHALRRTPSWSSLGGSEHQEMSFLEQENSSSWPSPAVTS
                     SSERIRGKRRAKALRWTRQKSVEEGEPPGQGEGPRSRPAAESTGLEATFPKTTPLAQA
                     DPAGVGTPPTGWDCLPSDCTASAAGSSTDDVELATEFPATEAWECELEGLLEERPALC
                     LSPQAPFPKLGWDDELRKPGAQIYMRFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFA
                     LVANGVRAAPLWDSKKQSFVGMLTITDFILVLHRYYRSPLVQIYEIEQHKIETWREIY
                     LQGCFKPLVSISPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIF
                     GSLLPRPSFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSALPVVNECGQV
                     VGLYSRFDVIHLAAQQTYNHLDMSVGEALRQRTLCLEGVLSCQPHESLGEVIDRIARE
                     QVHRLVLVDETQHLLGVVSLSDILQALVLSPAGIDALGA"
BASE COUNT      346 a    502 c    462 g    337 t
ORIGIN
```

Query Match 100.0%; Score 1647; DB 6; Length 1647;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1647; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 ttggtctggggctggccacatggagcccgggctggagcacgcactgcgcaggaccccttc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 TTGGTCTGGGGCTGGCCACATGGAGCCCGGGCTGGAGCACGCACTGCGCAGGACCCCTTC 60

Qy      61 ctggagcagccttgggggttctgagcatcaagagatgagcttcctagagcaagaaaacag 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CTGGAGCAGCCTTGGGGGTTCTGAGCATCAAGAGATGAGCTTCCTAGAGCAAGAAAACAG 120

Qy     121 cagctcatggccatcaccagctgtgaccagcagctcagaaagaatccgtgggaaacggag 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CAGCTCATGGCCATCACCAGCTGTGACCAGCAGCTCAGAAAGAATCCGTGGGAAACGGAG 180

Qy     181 ggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtca 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 GGCCAAAGCCTTGAGATGGACAAGGCAGAAGTCGGTGGAGGAAGGGGAGCCACCAGGTCA 240

Qy     241 gggggaaggtccccggtccaggccagctgctgagtccaccgggctggaggccacattccc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GGGGGAAGGTCCCCGGTCCAGGCCAGCTGCTGAGTCCACCGGGCTGGAGGCCACATTCCC 300

Qy     301 caagaccacaccccttggctcaagctgatcctgccggggtgggcactccaccaacagggtg 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 CAAGACCACACCCTTGGCTCAAGCTGATCCTGCCGGGGTGGGCACTCCACCAACAGGGTG 360

Qy     361 ggactgcctcccctctgactgtacagcctcagctgcaggctccagcacagatgatgtgga 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GGACTGCCTCCCCTCTGACTGTACAGCCTCAGCTGCAGGCTCCAGCACAGATGATGTGGA 420

Qy     421 gctggccacggagttcccagccacagaggcctgggagtgtgagctagaaggcctgctgga 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GCTGGCCACGGAGTTCCCAGCCACAGAGGCCTGGGAGTGTGAGCTAGAAGGCCTGCTGGA 480

Qy     481 agagaggcctgccctgtgcctgtccccgcaggccccatttcccaagctgggctgggatga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 AGAGAGGCCTGCCCTGTGCCTGTCCCCGCAGGCCCCATTTCCCAAGCTGGGCTGGGATGA 540

Qy     541 cgaactgcggaaacccggcgcccagatctacatgcgcttcatgcaggagcacacctgcta 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 CGAACTGCGGAAACCCGGCGCCCAGATCTACATGCGCTTCATGCAGGAGCACACCTGCTA 600

Qy     601 cgatgccatggcaactagctccaagctagtcatcttcgacaccatgctggagatcaagaa 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 CGATGCCATGGCAACTAGCTCCAAGCTAGTCATCTTCGACACCATGCTGGAGATCAAGAA 660

Qy     661 ggccttctttgctctggtggccaacggtgtgcgggcagcccctctatgggacagcaagaa 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GGCCTTCTTTGCTCTGGTGGCCAACGGTGTGCGGGCAGCCCCTCTATGGGACAGCAAGAA 720

Qy     721 gcagagctttgtggggatgctgaccatcactgacttcatcctggtgctgcatcgctacta 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 GCAGAGCTTTGTGGGGATGCTGACCATCACTGACTTCATCCTGGTGCTGCATCGCTACTA 780

Qy     781 caggtccccccctggtccagatctatgagattgaacaacataagattgagacctggaggga 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CAGGTCCCCCCCTGGTCCAGATCTATGAGATTGAACAACATAAGATTGAGACCTGGAGGGA 840

Qy     841 gatctacctgcaaggctgcttcaagcctctggtctccatctctcctaatgatagcctgtt 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 GATCTACCTGCAAGGCTGCTTCAAGCCTCTGGTCTCCATCTCTCCTAATGATAGCCTGTT 900

Qy     901 tgaagctgtctacaccctcatcaagaaccggatccatcgcctgcctgttcttgacccggt 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 TGAAGCTGTCTACACCCTCATCAAGAACCGGATCCATCGCCTGCCTGTTCTTGACCCGGT 960

Qy     961 gtcaggcaacgtactccacatcctcacacacaaacgcctgctcaagttcctgcacatctt 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GTCAGGCAACGTACTCCACATCCTCACACACAAACGCCTGCTCAAGTTCCTGCACATCTT 1020

Qy    1021 tggttccctgctgccccggccctccttcctctaccgcactatccaagatttgggcatcgg 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 TGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACTATCCAAGATTTGGGCATCGG 1080

Qy    1081 cacattccgagacttggctgtggtgctggagacagcacccatcctgactgcactggacat 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 CACATTCCGAGACTTGGCTGTGGTGCTGGAGACAGCACCCATCCTGACTGCACTGGACAT 1140

Qy    1141 ctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtggtcaggtcgtggg 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 CTTTGTGGACCGGCGTGTGTCTGCACTGCCTGTGGTCAACGAATGTGGTCAGGTCGTGGG 1200

Qy    1201 cctctattcccgctttgatgtgattcacctggctgcccagcaaacctacaaccacctgga 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 CCTCTATTCCCGCTTTGATGTGATTCACCTGGCTGCCCAGCAAACCTACAACCACCTGGA 1260

Qy    1261 catgagtgtgggagaagccctgaggcagaggacactatgtctggagggagtcctttcctg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 CATGAGTGTGGGAGAAGCCCTGAGGCAGAGGACACTATGTCTGGAGGGAGTCCTTTCCTG 1320

Qy    1321 ccagccccacgagagcttgggggaagtgatcgacaggattgctcgggagcaggtacacag 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 CCAGCCCCACGAGAGCTTGGGGGAAGTGATCGACAGGATTGCTCGGGAGCAGGTACACAG 1380

Qy    1381 gctggtgctagtggacgagacccagcatctcttgggcgtggtctccctctccgacatcct 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 GCTGGTGCTAGTGGACGAGACCCAGCATCTCTTGGGCGTGGTCTCCCTCTCCGACATCCT 1440

Qy    1441 tcaggcactggtgctcagccctgctggcatcgatgccctcggggcctgagaagatctgag 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 TCAGGCACTGGTGCTCAGCCCTGCTGGCATCGATGCCCTCGGGGCCTGAGAAGATCTGAG 1500

Qy    1501 tcctcaatcccaagccacctgcacacctggaagccaatgaagggaactggagaactcagc 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 TCCTCAATCCCAAGCCACCTGCACACCTGGAAGCCAATGAAGGGAACTGGAGAACTCAGC 1560

Qy    1561 cttcatcttccccccaccccccatttgctggttcagctatgattcaggtaggctctgccctg 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 CTTCATCTTCCCCCACCCCCCATTTGCTGGTTCAGCTATGATTCAGGTAGGCTCTGCCCTG 1620

Qy    1621 ggccatgacaccagcctcttagtcttc 1647
          |||||||||||||||||||||||||||
Db    1621 GGCCATGACACCAGCCTCTTAGTCTTC 1647
```

```
RESULT   2
HSA249977
LOCUS       HSA249977              2290 bp    mRNA    linear   PRI 07-APR-2000
DEFINITION  Homo sapiens mRNA for AMP-activated protein kinase gamma 3 subunit
            (AMPK gamma 3 gene).
ACCESSION   AJ249977
VERSION     AJ249977.1  GI:6688200
KEYWORDS    AMP-activated protein kinase; AMPK gamma 3 gene; gamma 3 subunit.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2290)
  AUTHORS   Cheung,P.C., Salt,I.P., Davies,S.P., Hardie,D.G. and Carling,D.
```

```
  TITLE     Characterization of AMP-activated protein kinase gamma-subunit
            isoforms and their role in AMP binding
  JOURNAL   Biochem. J. 346 Pt 3, 659-669 (2000)
  MEDLINE   20164049
REFERENCE   2  (bases 1 to 2290)
  AUTHORS   Carling,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-OCT-1999) Carling D., Cellular Stress Group, MRC
            Clinical Sciences Centre, Hammersmith Hospital, DuCane Road,
            London, W12 0NN, UNITED KINGDOM
FEATURES             Location/Qualifiers
     source          1. .2290
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            22. .1500
                     /gene="AMPK gamma 3"
     CDS             22. .1500
                     /gene="AMPK gamma 3"
                     /function="AMP-activated protein kinase regulatory
                     subunit"
                     /codon_start=1
                     /evidence=experimental
                     /product="AMP-activated protein kinase gamma 3 subunit"
                     /protein_id="CAB65117.1"
                     /db_xref="GI:6688201"
                     /translation="MEPGLEHALRRTPSWSSLGGSEHQEMSFLEQENSSSWPSPAVTS
                     SSERIRGKRRAKALRWTRQKSVEEGEPPGQGEGPRSRPAAESTGLEATFPKTTPLAQA
                     DPAGVGTPPTGWDCLPSDCTASAAGSSTDDVELATEFPATEAWECELEGLLEERPALC
                     LSPQAPFPKLGWDDELRKPGAQIYMRFIEEHTCYDAMATSSKLVIFDTMLEIKKAFFA
                     LVANGVRAAPLWDSKKQSFVGMLTITDFILVLHRYYRSPLVQIYEIEQHKIETWREIY
                     LQGCFKPLVSISPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIF
                     GSLLPRPSFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSALPVVNECGQV
                     VGLYSRFDVIHLAAQQTYNHLDMSVGEALRKRTLCLEGVLSCQPHESLGEVIDRIARE
                     QVHRLVLVDETQHLLGVVSLSDILQALVLSPAGIDPSGPEKI"
BASE COUNT      501 a    674 c    617 g    498 t
ORIGIN
```

Query Match 96.5%; Score 1590; DB 9; Length 2290;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1607; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

```
Qy       1 ttggtctggggctggccacatggagcccgggctggagcacgcactgcgcaggacccccttc 60
Db       3 TTGGTCTGGGGCTGGCCACATGGAGCCCGGGCTGGAGCACGCACTGCGCAGGACCCCTTC 62

Qy      61 ctggagcagccttgggggttctgagcatcaagagatgagcttcctagagcaagaaaacag 120
Db      63 CTGGAGCAGCCTTGGGGGTTCTGAGCATCAAGAGATGAGCTTCCTAGAGCAAGAAAACAG 122

Qy     121 cagctcatggccatcaccagctgtgaccagcagctcagaaagaatccgtgggaaacggag 180
Db     123 CAGCTCATGGCCATCACCAGCTGTGACCAGCAGCTCAGAAAGAATCCGTGGGAAACGGAG 182

Qy     181 ggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtca 240
Db     183 GGCCAAAGCCTTGAGATGGACAAGGCAGAAGTCGGTGGAGGAAGGGGAGCCACCAGGTCA 242

Qy     241 gggggaaggtccccggtccaggccagctgctgagtccaccgggctggaggccacattccc 300
Db     243 GGGGGAAGGTCCCCGGTCCAGGCCAGCTGCTGAGTCCACCGGGCTGGAGGCCACATTCCC 302

Qy     301 caagaccacaccccttggctcaagctgatcctgccggggtgggcactccaccaacagggtg 360
Db     303 CAAGACCACACCCTTGGCTCAAGCTGATCCTGCCGGGGTGGGCACTCCACCAACAGGGTG 362

Qy     361 ggactgcctccccctctgactgtacagcctcagctgcaggctccagcacagatgatgtgga 420
Db     363 GGACTGCCTCCCCTCTGACTGTACAGCCTCAGCTGCAGGCTCCAGCACAGATGATGTGGA 422

Qy     421 gctggccacggagttcccagccacagaggcctgggagtgtgagctagaaggcctgctgga 480
Db     423 GCTGGCCACGGAGTTCCCAGCCACAGAGGCCTGGGAGTGTGAGCTAGAAGGCCTGCTGGA 482

Qy     481 agagaggcctgccctgtgcctgtccccgcaggccccatttcccaagctgggctgggatga 540
Db     483 AGAGAGGCCTGCCCTGTGCCTGTCCCCGCAGGCCCCATTTCCCAAGCTGGGCTGGGATGA 542

Qy     541 cgaactgcggaaaccccggcgcccagatctacatgcgcttcatgcaggagcacacctgcta 600
Db     543 CGAACTGCGGAAACCCGGCGCCCAGATCTACATGCGCTTCATCGAGGAGCACACCTGCTA 602

Qy     601 cgatgccatggcaactagctccaagctagtcatcttcgacaccatgctggagatcaagaa 660
Db     603 CGATGCCATGGCAACTAGCTCCAAGCTAGTCATCTTCGACACCATGCTGGAGATCAAGAA 662

Qy     661 ggccttctttgctctggtggccaacggtgtgcgggcagcccctctatgggacagcaagaa 720
Db     663 GGCCTTCTTTGCTCTGGTGGCCAACGGTGTGCGGGCAGCCCCTCTATGGGACAGCAAGAA 722

Qy     721 gcagagctttgtggggatgctgaccatcactgacttcatcctggtgctgcatcgctacta 780
Db     723 GCAGAGCTTTGTGGGGATGCTGACCATCACTGACTTCATCCTGGTGCTGCATCGCTACTA 782

Qy     781 caggtccccccctggtccagatctatgagattgaacaacataagattgagacctggaggga 840
Db     783 CAGGTCCCCCCTGGTCCAGATCTATGAGATTGAACAACATAAGATTGAGACCTGGAGGGA 842

Qy     841 gatctacctgcaaggctgcttcaagcctctggtctccatctctcctaatgatagcctgtt 900
Db     843 GATCTACCTGCAAGGCTGCTTCAAGCCTCTGGTCTCCATCTCTCCTAATGATAGCCTGTT 902

Qy     901 tgaagctgtctacaccctcatcaagaaccggatccatcgcctgcctgttcttgacccggt 960
Db     903 TGAAGCTGTCTACACCCTCATCAAGAACCGGATCCATCGCCTGCCTGTTCTTGACCCGGT 962

Qy     961 gtcaggcaacgtactccacatcctcacacacaaacgcctgctcaagttcctgcacatctt 1020
Db     963 GTCAGGCAACGTACTCCACATCCTCACACACAAACGCCTGCTCAAGTTCCTGCACATCTT 1022

Qy    1021 tggttccctgctgccccggccctccttcctctaccgcactatccaagatttgggcatcgg 1080
Db    1023 TGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACTATCCAAGATTTGGGCATCGG 1082

Qy    1081 cacattccgagacttggctgtggtgctggagacagcacccatcctgactgcactggacat 1140
Db    1083 CACATTCCGAGACTTGGCTGTGGTGCTGGAGACAGCACCCATCCTGACTGCACTGGACAT 1142

Qy    1141 ctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtggtcaggtcgtggg 1200
Db    1143 CTTTGTGGACCGGCGTGTGTCTGCACTGCCTGTGGTCAACGAATGTGGTCAGGTCGTGGG 1202

Qy    1201 cctctattcccgctttgatgtgattcacctggctgcccagcaaacctacaaccacctgga 1260
Db    1203 CCTCTATTCCCGCTTTGATGTGATTCACCTGGCTGCCCAGCAAACCTACAACCACCTGGA 1262

Qy    1261 catgagtgtgggagaagccctgaggcagaggacactatgtctggagggagtcctttcctg 1320
Db    1263 CATGAGTGTGGGAGAAGCCCTGAGGAAGAGGACACTATGTCTGGAGGGAGTCCTTTCCTG 1322

Qy    1321 ccagccccacgagagcttgggggaagtgatcgacaggattgctcgggagcaggtacacag 1380
Db    1323 CCAGCCCCACGAGAGCTTGGGGGAAGTGATCGACAGGATTGCTCGGGAGCAGGTACACAG 1382

Qy    1381 gctggtgctagtggacgagacccagcatctcttgggcgtggtctccctctccgacatcct 1440
Db    1383 GCTGGTGCTAGTGGACGAGACCCAGCATCTCTTGGGCGTGGTCTCCCTCTCCGACATCCT 1442

Qy    1441 tcaggcactggtgctcagccctgctggcatcgatgccctcggggcctgagaagatctgag 1500
Db    1443 TCAGGCACTGGTGCTCAGCCCTGCTGGCATCGAT-CCCTCGGGGCCTGAGAAGATCTGAG 1501

Qy    1501 tcctcaatcccaagccacctgcacacctggaagccaatgaagggaactggagaactcagc 1560
Db    1502 TCCTCAATCCCAAGCCACCTGCACACCTGGAAGCCAATGAAGGGAACTGGAGAACTCAGC 1561
```

```
Qy   1561 cttcatcttcccccaccccccatttgctggttcagctatgattcaggtaggctctgccc 1618
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1562 CTTCATCTTCCCCCACCCCCATTTGCTGGTTCAGCTATGATTCAGGCTTCTTGAGGCC 1619


RESULT   3
AX099802
LOCUS       AX099802.                2115 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 29 from Patent WO0120003.
ACCESSION   AX099802
VERSION     AX099802.1  GI:13538836
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2115)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 29 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .2115
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             1. .1395
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35801.1"
                     /db_xref="GI:13538837"
                     /translation="MSFLEQENSSSWPSPAVTSSSERIRGKRRAKALRWTRQKSVEEG
                     EPPGQGEGPRSRPTAESTGLEATFPKTTPLAQADPAGVGTPPTGWDCLPSDCTASAAG
                     SSTDDVELATEFPATEAWECELEGLLEERPALCLSPQAPFPKLGWDDELRKPGAQIYM
                     RFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEQHKIETWREIYLQGCFKPLVSISPNDSLFEAVYTLI
                     KNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNECGQVVGLYSRFDVIHLAAQQTYNHLDMSV
                     GEALRQRTLCLEGVLSCQPHESLGEVIDRIAREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT      460 a    622 c    562 g    471 t
ORIGIN


  Query Match           88.2%;  Score 1453;  DB 6;  Length 2115;
  Best Local Similarity 97.3%;  Pred. No. 0;
  Matches 1501;  Conservative    0;  Mismatches   35;  Indels    7;  Gaps    2;

Qy     95 atgagcttcctagagcaagaaaacagcagctcatggccatcaccagctgtgaccagcagc 154
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAGCTTCCTAGAGCAAGAAAACAGCAGCTCATGGCCATCACCAGCTGTGACCAGCAGC 60

Qy    155 tcagaaagaatccgtgggaaacggagggccaaagccttgagatggacaaggcagaagtcg 214
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TCAGAAAGAATCCGTGGGAAACGGAGGGCCAAAGCCTTGAGATGGACAAGGCAGAAGTCG 120

Qy    215 gtggaggaaggggagccaccaggtcaggggaaggtccccggtccaggccagctgctgag 274
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGGAGGAAGGGGAGCCACCAGGTCAGGGGGAAGGTCCCCGGTCCAGGCCAACTGCTGAG 180

Qy    275 tccaccgggctggaggccacattccccaagaccacaccttggctcaagctgatcctgcc 334
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TCCACCGGGCTGGAGGCCACATTCCCCAAGACCACACCCTTGGCTCAAGCTGATCCTGCC 240

Qy    335 ggggtgggcactccaccaacagggtgggactgcctcccctctgactgtacagcctcagct 394
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGGGTGGGCACTCCACCAACAGGGTGGGACTGCCTCCCCTCTGACTGTACAGCCTCAGCT 300

Qy    395 gcaggctccagcacagatgatgtggagctggccacggagttcccagccacagaggcctgg 454
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GCAGGCTCCAGCACAGATGATGTGGAGCTGGCCACGGAGTTCCCAGCCACAGAGGCCTGG 360

Qy    455 gagtgtgagctagaaggcctgctggaagagaggcctgccctgtgcctgtccccgcaggcc 514
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GAGTGTGAGCTAGAAGGCCTGCTGGAAGAGAGGCCTGCCCTGTGCCTGTCCCCGCAGGCC 420

Qy    515 ccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcccagatctacatg 574
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CCATTTCCCAAGCTGGGCTGGGATGACGAACTGCGGAAACCCGGCGCCCAGATCTACATG 480

Qy    575 cgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagctagtcatc 634
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CGCTTCATGCAGGAGCACACCTGCTACGATGCCATGGCAACTAGCTCCAAGCTAGTCATC 540

Qy    635 ttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacggtgtgcgg 694
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TTCGACACCATGCTGGAGATCAAGAAGGCCTTCTTTGCTCTGGTGGCCAACGGTGTGCGG 600

Qy    695 gcagcccctctatgggacagcaagaagcagagctttgtggggatgctgaccatcactgac 754
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCAGCCCCTCTATGGGACAGCAAGAAGCAGAGCTTTGTGGGGATGCTGACCATCACTGAC 660

Qy    755 ttcatcctggtgctgcatcgctactacaggtccccctggtccagatctatgagattgaa 814
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TTCATCCTGGTGCTGCATCGCTACTACAGGTCCCCCCTGGTCCAGATCTATGAGATTGAA 720

Qy    815 caacataagattgagacctggagggagatctacctgcaaggctgcttcaagcctctggtc 874
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CAACATAAGATTGAGACCTGGAGGGAGATCTACCTGCAAGGCTGCTTCAAGCCTCTGGTC 780

Qy    875 tccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaagaaccggatc 934
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TCCATCTCTCCTAATGATAGCCTGTTTGAAGCTGTCTACACCCTCATCAAGAACCGGATC 840

Qy    935 catcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctcacacacaaa 994
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CATCGCCTGCCTGTTCTTGACCCGGTGTCAGGCAACGTACTCCACATCCTCACACACAAA 900

Qy    995 cgcctgctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctac 1054
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CGCCTGCTCAAGTTCCTGCACATCTTTGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTAC 960

Qy   1055 cgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgctggagaca 1114
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CGCACTATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCTGTGGTGCTGGAGACA 1020

Qy   1115 gcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtg 1174
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GCACCCATCCTGACTGCACTGGACATCTTTGTGGACCGGCGTGTGTCTGCACTGCCTGTG 1080

Qy   1175 gtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattcacctggct 1234
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GTCAACGAATGTGGTCAGGTCGTGGGCCTCTATTCCCGCTTTGATGTGATTCACCTGGCT 1140

Qy   1235 gcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggcagaggaca 1294
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GCCCAGCAAACCTACAACCACCTGGACATGAGTGTGGGAGAAGCCCTGAGGCAGAGGACA 1200

Qy   1295 ctatgtctggagggagtcctttcctgccagccccacgagagcttgggggaagtgatcgac 1354
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTATGTCTGGAGGGAGTCCTTTCCTGCCAGCCCCACGAGAGCTTGGGGGAAGTGATCGAC 1260

Qy   1355 aggattgctcgggagcaggtacacaggctggtgctagtggacgagacccagcatctcttg 1414
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AGGATTGCTCGGGAGCAGGTACACAGGCTGGTGCTAGTGGACGAGACCCAGCATCTCTTG 1320

Qy   1415 ggcgtggtctccctctccgacatccttcaggcactggtgctcagccctgctggcatcgat 1474
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GGCGTGGTCTCCCTCTCCGACATCCTTCAGGCACTGGTGCTCAGCCCTGCTGGCATCGAT 1380

Qy   1475 gccctcggggcctgagaagatctgagtcctcaatcccaagccacctgcacacctggaagc 1534
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GCCCTCGGGGCCTGAGAAGATCTGAGTCCTCAATCCCAAGCCAACTGCACA-CTGGAAGC 1439
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 14:46:57 ; Search time 7316.32 Seconds
(without alignments)
4710.839 Million cell updates/sec

Title: US-09-826-581-5
Perfect score: 1647
Sequence: 1 ttggtctggggctggccaca.....acaccagcctcttagtcttc 1647

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1647	100.0	1647	6	AX281582	AX281582 Sequence
2	1590	96.5	2290	9	HSA249977	AJ249977 Homo sapi
3	1453	88.2	2115	6	AX099802	AX099802 Sequence
4	1453	88.2	2115	9	AF214519	AF214519 Homo sapi
5	1447	87.9	2109	6	AX099776	AX099776 Sequence
6	1172.8	71.2	2022	6	AX099804	AX099804 Sequence
7	1140.4	69.2	1873	4	AF214520	AF214520 Sus scrof
8	1140.4	69.2	1873	6	AX099800	AX099800 Sequence
9	1134.4	68.9	1867	6	AX099774	AX099774 Sequence
10	414.8	25.2	1328	10	RNAMPKGAM	X95578 R.norvegicu
11	414.8	25.2	1550	10	RNU42413	U42413 Rattus norv
12	407.2	24.7	989	6	AX281579	AX281579 Sequence
c 13	407.2	24.7	196554	2	AC073128	AC073128 Homo sapi
14	396.8	24.1	1623	10	AF036535	AF036535 Mus muscu
15	395.4	24.0	3132	10	BC015283	BC015283 Mus muscu
c 16	395.2	24.0	206854	9	AC009974	AC009974 Homo sapi
17	393.4	23.9	1578	9	HSU42412	U42412 Human 5'-AM
18	393.4	23.9	1677	9	BC000358	BC000358 Homo sapi
19	366.6	22.3	1167	9	AB025580	AB025580 Homo sapi
20	366.6	22.3	1435	6	AR139104	AR139104 Sequence
21	366.6	22.3	2194	9	AF087875	AF087875 Homo sapi
22	366.6	22.3	2203	9	BC020540	BC020540 Homo sapi
23	366.6	22.3	2223	9	AK001887	AK001887 Homo sapi
24	363.4	22.1	2062	9	HSA249976	AJ249976 Homo sapi
25	300.8	18.3	2082	3	AF094764	AF094764 Drosophil
26	299.2	18.2	4666	3	AY070541	AY070541 Drosophil
27	278.8	16.9	1014	6	AX281581	AX281581 Sequence
28	278.8	16.9	152129	2	AC027416	AC027416 Homo sapi
29	278.8	16.9	196554	2	AC073128	AC073128 Homo sapi
30	273.6	16.6	5888	4	AF214521	AF214521 Sus scrof
c 31	267.4	16.2	227724	2	AF336381	AF336381 Mus muscu
c 32	198.2	12.0	152129	2	AC027416	AC027416 Homo sapi
33	168.2	10.2	1722	6	AX281580	AX281580 Sequence
34	131	8.0	3497	3	AF094763	AF094763 Drosophil
35	127.2	7.7	80069	2	AC019671	AC019671 Drosophil
36	127.2	7.7	173634	3	AC009344	AC009344 Drosophil
37	127.2	7.7	195868	3	AC008308	AC008308 Drosophil
38	127.2	7.7	230266	3	AE003733	AE003733 Drosophil
39	91.4	5.5	106	11	G67375	G67375 A898 Human
40	84.8	5.1	151	6	AX305717	AX305717 Sequence
41	83	5.0	142903	2	AL627254	AL627254 Danio rer
c 42	78.4	4.8	110000	2	CEY111B2_1	Continuation (2 of
c 43	78.4	4.8	110000	2	CEY111B2_2	Continuation (3 of
c 44	78.4	4.8	257702	3	CEY111B2A	AL132904 Caenorhab
45	75	4.6	14411	4	AF329081	AF329081 Bos tauru

ALIGNMENTS

RESULT 1
AX281582
LOCUS AX281582 1647 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 5 from Patent WO0177305.
ACCESSION AX281582
VERSION AX281582.1 GI:16608833
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Andersson,L., Luthman,H. and Marklund,S.
TITLE Variants of the human amp-activated protein kinase gamma 3 subunit
JOURNAL Patent: WO 0177305-A 5 18-OCT-2001;
Arexis AB (SE)
FEATURES Location/Qualifiers
source 1. .1647
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS 20. .1489
/note="unnamed protein product"
/codon_start=1

```
                     /protein_id="CAD10589.1"
                     /db_xref="GI:16608834"
                     /translation="MEPGLEHALRRTPSWSSLGGSEHQEMSFLEQENSSSWPSPAVTS
                     SSERIRGKRRAKALRWTRQKSVEEGEPPGQGEGPRSRPAAESTGLEATFPKTTPLAQA
                     DPAGVGTPPTGWDCLPSDCTASAAGSSTDDVELATEFPATEAWECELEGLLEERPALC
                     LSPQAPFPKLGWDDELRKPGAQIYMRFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFA
                     LVANGVRAAPLWDSKKQSFVGMLTITDFILVLHRYYRSPLVQIYEIEQHKIETWREIY
                     LQGCFKPLVSISPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIF
                     GSLLPRPSFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSALPVVNECGQV
                     VGLYSRFDVIHLAAQQTYNHLDMSVGEALRQRTLCLEGVLSCQPHESLGEVIDRIARE
                     QVHRLVLVDETQHLLGVVSLSDILQALVLSPAGIDALGA"
BASE COUNT      346 a    502 c    462 g    337 t
ORIGIN
```

Query Match 100.0%; Score 1647; DB 6; Length 1647;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1647; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 ttggtctggggctggccacatggagcccgggctggagcacgcactgcgcaggacccccttc 60
Db       1 TTGGTCTGGGGCTGGCCACATGGAGCCCGGGCTGGAGCACGCACTGCGCAGGACCCCTTC 60

Qy      61 ctggagcagccttgggggttctgagcatcaagagatgagcttcctagagcaagaaaacag 120
Db      61 CTGGAGCAGCCTTGGGGGTTCTGAGCATCAAGAGATGAGCTTCCTAGAGCAAGAAAACAG 120

Qy     121 cagctcatggccatcaccagctgtgaccagcagctcagaaagaatccgtgggaaacggag 180
Db     121 CAGCTCATGGCCATCACCAGCTGTGACCAGCAGCTCAGAAAGAATCCGTGGGAAACGGAG 180

Qy     181 ggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtca 240
Db     181 GGCCAAAGCCTTGAGATGGACAAGGCAGAAGTCGGTGGAGGAAGGGGAGCCACCAGGTCA 240

Qy     241 gggggaaggtccccggtccaggccagctgctgagtccaccgggctggaggccacattccc 300
Db     241 GGGGGAAGGTCCCCGGTCCAGGCCAGCTGCTGAGTCCACCGGGCTGGAGGCCACATTCCC 300

Qy     301 caagaccacacccttggctcaagctgatcctgccggggtgggcactccaccaacagggtg 360
Db     301 CAAGACCACACCCTTGGCTCAAGCTGATCCTGCCGGGGTGGGCACTCCACCAACAGGGTG 360

Qy     361 ggactgcctcccctctgactgtacagcctcagctgcaggctccagcacagatgatgtgga 420
Db     361 GGACTGCCTCCCCTCTGACTGTACAGCCTCAGCTGCAGGCTCCAGCACAGATGATGTGGA 420

Qy     421 gctggccacggagttcccagccacagaggcctgggagtgtgagctagaaggcctgctgga 480
Db     421 GCTGGCCACGGAGTTCCCAGCCACAGAGGCCTGGGAGTGTGAGCTAGAAGGCCTGCTGGA 480

Qy     481 agagaggcctgccctgtgcctgtccccgcaggccccatttcccaagctgggctgggatga 540
Db     481 AGAGAGGCCTGCCCTGTGCCTGTCCCCGCAGGCCCCATTTCCCAAGCTGGGCTGGGATGA 540

Qy     541 cgaactgcggaaacccggcgcccagatctacatgcgcttcatgcaggagcacacctgcta 600
Db     541 CGAACTGCGGAAACCCGGCGCCCAGATCTACATGCGCTTCATGCAGGAGCACACCTGCTA 600

Qy     601 cgatgccatggcaactagctccaagctagtcatcttcgacaccatgctggagatcaagaa 660
Db     601 CGATGCCATGGCAACTAGCTCCAAGCTAGTCATCTTCGACACCATGCTGGAGATCAAGAA 660

Qy     661 ggccttctttgctctggtggccaacggtgtgcgggcagcccctctatgggacagcaagaa 720
Db     661 GGCCTTCTTTGCTCTGGTGGCCAACGGTGTGCGGGCAGCCCCTCTATGGGACAGCAAGAA 720

Qy     721 gcagagctttgtggggatgctgaccatcactgacttcatcctggtgctgcatcgctacta 780
Db     721 GCAGAGCTTTGTGGGGATGCTGACCATCACTGACTTCATCCTGGTGCTGCATCGCTACTA 780

Qy     781 caggtccccccctggtccagatctatgagattgaacaacataagattgagacctggaggga 840
Db     781 CAGGTCCCCCCTGGTCCAGATCTATGAGATTGAACAACATAAGATTGAGACCTGGAGGGA 840

Qy     841 gatctacctgcaaggctgcttcaagcctctggtctccatctctcctaatgatagcctgtt 900
Db     841 GATCTACCTGCAAGGCTGCTTCAAGCCTCTGGTCTCCATCTCTCCTAATGATAGCCTGTT 900

Qy     901 tgaagctgtctacaccctcatcaagaaccggatccatcgcctgcctgttcttgacccggt 960
Db     901 TGAAGCTGTCTACACCCTCATCAAGAACCGGATCCATCGCCTGCCTGTTCTTGACCCGGT 960

Qy     961 gtcaggcaacgtactccacatcctcacacacaaacgcctgctcaagttcctgcacatctt 1020
Db     961 GTCAGGCAACGTACTCCACATCCTCACACACAAACGCCTGCTCAAGTTCCTGCACATCTT 1020

Qy    1021 tggttccctgctgccccggccctccttcctctaccgcactatccaagatttgggcatcgg 1080
Db    1021 TGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACTATCCAAGATTTGGGCATCGG 1080

Qy    1081 cacattccgagacttggctgtggtgctggagacagcacccatcctgactgcactggacat 1140
Db    1081 CACATTCCGAGACTTGGCTGTGGTGCTGGAGACAGCACCCATCCTGACTGCACTGGACAT 1140

Qy    1141 ctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtggtcaggtcgtggg 1200
Db    1141 CTTTGTGGACCGGCGTGTGTCTGCACTGCCTGTGGTCAACGAATGTGGTCAGGTCGTGGG 1200

Qy    1201 cctctattcccgctttgatgtgattcacctggctgcccagcaaacctacaaccacctgga 1260
Db    1201 CCTCTATTCCCGCTTTGATGTGATTCACCTGGCTGCCCAGCAAACCTACAACCACCTGGA 1260

Qy    1261 catgagtgtgggagaagccctgaggcagaggacactatgtctggagggagtcctttcctg 1320
Db    1261 CATGAGTGTGGGAGAAGCCCTGAGGCAGAGGACACTATGTCTGGAGGGAGTCCTTTCCTG 1320

Qy    1321 ccagccccacgagagcttgggggaagtgatcgacaggattgctcgggagcaggtacacag 1380
Db    1321 CCAGCCCCACGAGAGCTTGGGGGAAGTGATCGACAGGATTGCTCGGGAGCAGGTACACAG 1380

Qy    1381 gctggtgctagtggacgagacccagcatctcttgggcgtggtctccctctccgacatcct 1440
Db    1381 GCTGGTGCTAGTGGACGAGACCCAGCATCTCTTGGGCGTGGTCTCCCTCTCCGACATCCT 1440

Qy    1441 tcaggcactggtgctcagccctgctggcatcgatgccctcggggcctgagaagatctgag 1500
Db    1441 TCAGGCACTGGTGCTCAGCCCTGCTGGCATCGATGCCCTCGGGGCCTGAGAAGATCTGAG 1500

Qy    1501 tcctcaatcccaagccacctgcacacctggaagccaatgaagggaactggagaactcagc 1560
Db    1501 TCCTCAATCCCAAGCCACCTGCACACCTGGAAGCCAATGAAGGGAACTGGAGAACTCAGC 1560

Qy    1561 cttcatcttccccacccccatttgctggttcagctatgattcaggtaggctctgccctg 1620
Db    1561 CTTCATCTTCCCCCACCCCCATTTGCTGGTTCAGCTATGATTCAGGTAGGCTCTGCCCTG 1620

Qy    1621 ggccatgacaccagcctcttagtcttc 1647
Db    1621 GGCCATGACACCAGCCTCTTAGTCTTC 1647
```

```
RESULT   2
HSA249977
LOCUS       HSA249977              2290 bp    mRNA    linear   PRI 07-APR-2000
DEFINITION  Homo sapiens mRNA for AMP-activated protein kinase gamma 3 subunit
            (AMPK gamma 3 gene).
ACCESSION   AJ249977
VERSION     AJ249977.1  GI:6688200
KEYWORDS    AMP-activated protein kinase; AMPK gamma 3 gene; gamma 3 subunit.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2290)
  AUTHORS   Cheung,P.C., Salt,I.P., Davies,S.P., Hardie,D.G. and Carling,D.
```

```
  TITLE     Characterization of AMP-activated protein kinase gamma-subunit
            isoforms and their role in AMP binding
  JOURNAL   Biochem. J. 346 Pt 3, 659-669 (2000)
  MEDLINE   20164049
REFERENCE   2  (bases 1 to 2290)
  AUTHORS   Carling,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-OCT-1999) Carling D., Cellular Stress Group, MRC
            Clinical Sciences Centre, Hammersmith Hospital, DuCane Road,
            London, W12 0NN, UNITED KINGDOM
FEATURES             Location/Qualifiers
     source          1. .2290
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            22. .1500
                     /gene="AMPK gamma 3"
     CDS             22. .1500
                     /gene="AMPK gamma 3"
                     /function="AMP-activated protein kinase regulatory
                     subunit"
                     /codon_start=1
                     /evidence=experimental
                     /product="AMP-activated protein kinase gamma 3 subunit"
                     /protein_id="CAB65117.1"
                     /db_xref="GI:6688201"
                     /translation="MEPGLEHALRRTPSWSSLGGSEHQEMSFLEQENSSSWPSPAVTS
                     SSERIRGKRRAKALRWTRQKSVEEGEPPGQGEGPRSRPAAESTGLEATFPKTTPLAQA
                     DPAGVGTPPTGWDCLPSDCTASAAGSSTDDVELATEFPATEAWECELEGLLEERPALC
                     LSPQAPFPKLGWDDELRKPGAQIYMRFIEEHTCYDAMATSSKLVIFDTMLEIKKAFFA
                     LVANGVRAAPLWDSKKQSFVGMLTITDFILVLHRYYRSPLVQIYEIEQHKIETWREIY
                     LQGCFKPLVSISPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIF
                     GSLLPRPSFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSALPVVNECGQV
                     VGLYSRFDVIHLAAQQTYNHLDMSVGEALRKRTLCLEGVLSCQPHESLGEVIDRIARE
                     QVHRLVLVDETQHLLGVVSLSDILQALVLSPAGIDPSGPEKI"
BASE COUNT      501 a    674 c    617 g    498 t
ORIGIN
```

```
  Query Match             96.5%;  Score 1590;  DB 9;  Length 2290;
  Best Local Similarity   99.3%;  Pred. No. 0;
  Matches 1607;  Conservative    0;  Mismatches   10;  Indels    1;  Gaps    1;

Qy       1 ttggtctggggctggccacatggagcccgggctggagcacgcactgcgcaggaccccttc 60
Db       3 TTGGTCTGGGGCTGGCCACATGGAGCCCGGGCTGGAGCACGCACTGCGCAGGACCCCTTC 62

Qy      61 ctggagcagccttgggggttctgagcatcaagagatgagcttcctagagcaagaaaacag 120
Db      63 CTGGAGCAGCCTTGGGGGTTCTGAGCATCAAGAGATGAGCTTCCTAGAGCAAGAAAACAG 122

Qy     121 cagctcatggccatcaccagctgtgaccagcagctcagaaagaatccgtgggaaacggag 180
Db     123 CAGCTCATGGCCATCACCAGCTGTGACCAGCAGCTCAGAAAGAATCCGTGGGAAACGGAG 182

Qy     181 ggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtca 240
Db     183 GGCCAAAGCCTTGAGATGGACAAGGCAGAAGTCGGTGGAGGAAGGGGAGCCACCAGGTCA 242

Qy     241 gggggaaggtccccggtccaggccagctgctgagtccaccgggctggaggccacattccc 300
Db     243 GGGGGAAGGTCCCCGGTCCAGGCCAGCTGCTGAGTCCACCGGGCTGGAGGCCACATTCCC 302

Qy     301 caagaccacacccttggctcaagctgatcctgccggggtgggcactccaccaacagggtg 360
Db     303 CAAGACCACACCCTTGGCTCAAGCTGATCCTGCCGGGGTGGGCACTCCACCAACAGGGTG 362

Qy     361 ggactgcctccccctctgactgtacagcctcagctgcaggctccagcacagatgatgtgga 420
Db     363 GGACTGCCTCCCCTCTGACTGTACAGCCTCAGCTGCAGGCTCCAGCACAGATGATGTGGA 422

Qy     421 gctggccacggagttcccagccacagaggcctgggagtgtgagctagaaggcctgctgga 480
Db     423 GCTGGCCACGGAGTTCCCAGCCACAGAGGCCTGGGAGTGTGAGCTAGAAGGCCTGCTGGA 482

Qy     481 agagaggcctgccctgtgcctgtccccgcaggccccatttcccaagctgggctgggatga 540
Db     483 AGAGAGGCCTGCCCTGTGCCTGTCCCCGCAGGCCCCATTTCCCAAGCTGGGCTGGGATGA 542

Qy     541 cgaactgcggaaacccggcgcccagatctacatgcgcttcatgcaggagcacacctgcta 600
Db     543 CGAACTGCGGAAACCCGGCGCCCAGATCTACATGCGCTTCATCGAGGAGCACACCTGCTA 602

Qy     601 cgatgccatggcaactagctccaagctagtcatcttcgacaccatgctggagatcaagaa 660
Db     603 CGATGCCATGGCAACTAGCTCCAAGCTAGTCATCTTCGACACCATGCTGGAGATCAAGAA 662

Qy     661 ggccttctttgctctggtggccaacggtgtgcgggcagcccctctatgggacagcaagaa 720
Db     663 GGCCTTCTTTGCTCTGGTGGCCAACGGTGTGCGGGCAGCCCCTCTATGGGACAGCAAGAA 722

Qy     721 gcagagctttgtggggatgctgaccatcactgacttcatcctggtgctgcatcgctacta 780
Db     723 GCAGAGCTTTGTGGGGATGCTGACCATCACTGACTTCATCCTGGTGCTGCATCGCTACTA 782

Qy     781 caggtccccccctggtccagatctatgagattgaacaacataagattgagacctggaggga 840
Db     783 CAGGTCCCCCCTGGTCCAGATCTATGAGATTGAACAACATAAGATTGAGACCTGGAGGGA 842

Qy     841 gatctacctgcaaggctgcttcaagcctctggtctccatctctcctaatgatagcctgtt 900
Db     843 GATCTACCTGCAAGGCTGCTTCAAGCCTCTGGTCTCCATCTCTCCTAATGATAGCCTGTT 902

Qy     901 tgaagctgtctacaccctcatcaagaaccggatccatcgcctgcctgttcttgacccggt 960
Db     903 TGAAGCTGTCTACACCCTCATCAAGAACCGGATCCATCGCCTGCCTGTTCTTGACCCGGT 962

Qy     961 gtcaggcaacgtactccacatcctcacacacaaacgcctgctcaagttcctgcacatctt 1020
Db     963 GTCAGGCAACGTACTCCACATCCTCACACACAAACGCCTGCTCAAGTTCCTGCACATCTT 1022

Qy    1021 tggttccctgctgccccggccctccttcctctaccgcactatccaagatttgggcatcgg 1080
Db    1023 TGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACTATCCAAGATTTGGGCATCGG 1082

Qy    1081 cacattccgagacttggctgtggtgctggagacagcacccatcctgactgcactggacat 1140
Db    1083 CACATTCCGAGACTTGGCTGTGGTGCTGGAGACAGCACCCATCCTGACTGCACTGGACAT 1142

Qy    1141 ctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtggtcaggtcgtggg 1200
Db    1143 CTTTGTGGACCGGCGTGTGTCTGCACTGCCTGTGGTCAACGAATGTGGTCAGGTCGTGGG 1202

Qy    1201 cctctattcccgctttgatgtgattcacctggctgcccagcaaacctacaaccacctgga 1260
Db    1203 CCTCTATTCCCGCTTTGATGTGATTCACCTGGCTGCCCAGCAAACCTACAACCACCTGGA 1262

Qy    1261 catgagtgtgggagaagccctgaggcagaggacactatgtctggagggagtcctttcctg 1320
Db    1263 CATGAGTGTGGGAGAAGCCCTGAGGAAGAGGACACTATGTCTGGAGGGAGTCCTTTCCTG 1322

Qy    1321 ccagccccacgagagcttgggggaagtgatcgacaggattgctcgggagcaggtacacag 1380
Db    1323 CCAGCCCCACGAGAGCTTGGGGGAAGTGATCGACAGGATTGCTCGGGAGCAGGTACACAG 1382

Qy    1381 gctggtgctagtggacgagacccagcatctcttgggcgtggtctccctctccgacatcct 1440
Db    1383 GCTGGTGCTAGTGGACGAGACCCAGCATCTCTTGGGCGTGGTCTCCCTCTCCGACATCCT 1442

Qy    1441 tcaggcactggtgctcagccctgctggcatcgatgccctcggggcctgagaagatctgag 1500
Db    1443 TCAGGCACTGGTGCTCAGCCCTGCTGGCATCGAT-CCCTCGGGGCCTGAGAAGATCTGAG 1501

Qy    1501 tcctcaatcccaagccacctgcacacctggaagccaatgaagggaactggagaactcagc 1560
Db    1502 TCCTCAATCCCAAGCCACCTGCACACCTGGAAGCCAATGAAGGGAACTGGAGAACTCAGC 1561
```

```
Qy   1561 cttcatcttccccacccccatttgctggttcagctatgattcaggtaggctctgccc 1618
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1562 CTTCATCTTCCCCCACCCCCATTTGCTGGTTCAGCTATGATTCAGGCTTCTTCAGGCC 1619


RESULT   3
AX099802
LOCUS       AX099802                2115 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 29 from Patent WO0120003.
ACCESSION   AX099802
VERSION     AX099802.1  GI:13538836
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2115)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 29 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .2115
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             1. .1395
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35801.1"
                     /db_xref="GI:13538837"
                     /translation="MSFLEQENSSSWPSPAVTSSSERIRGKRRAKALRWTRQKSVEEG
                     EPPGQGEGPRSRPTAESTGLEATFPKTTPLAQADPAGVGTPPTGWDCLPSDCTASAAG
                     SSTDDVELATEFPATEAWECELEGLLEERPALCLSPQAPFPKLGWDDELRKPGAQIYM
                     RFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEQHKIETWREIYLQGCFKPLVSISPNDSLFEAVYTLI
                     KNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNECGQVVGLYSRFDVIHLAAQQTYNHLDMSV
                     GEALRQRTLCLEGVLSCQPHESLGEVIDRIAREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT      460 a    622 c    562 g    471 t
ORIGIN


  Query Match           88.2%;  Score 1453;  DB 6;  Length 2115;
  Best Local Similarity  97.3%;  Pred. No. 0;
  Matches 1501;  Conservative    0;  Mismatches   35;  Indels    7;  Gaps    2;

Qy     95 atgagcttcctagagcaagaaaacagcagctcatggccatcaccagctgtgaccagcagc 154
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAGCTTCCTAGAGCAAGAAAACAGCAGCTCATGGCCATCACCAGCTGTGACCAGCAGC 60

Qy    155 tcagaaagaatccgtgggaaacggagggccaaagccttgagatggacaaggcagaagtcg 214
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TCAGAAAGAATCCGTGGGAAACGGAGGGCCAAAGCCTTGAGATGGACAAGGCAGAAGTCG 120

Qy    215 gtggaggaaggggagccaccaggtcaggggggaaggtccccggtccaggccagctgctgag 274
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 GTGGAGGAAGGGGAGCCACCAGGTCAGGGGGAAGGTCCCCGGTCCAGGCCAACTGCTGAG 180

Qy    275 tccaccgggctggaggccacattccccaagaccacacccttggctcaagctgatcctgcc 334
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TCCACCGGGCTGGAGGCCACATTCCCCAAGACCACACCCTTGGCTCAAGCTGATCCTGCC 240

Qy    335 ggggtgggcactccaccaacagggtgggactgcctcccctctgactgtacagcctcagct 394
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGGGTGGGCACTCCACCAACAGGGTGGGACTGCCTCCCCTCTGACTGTACAGCCTCAGCT 300

Qy    395 gcaggctccagcacagatgatgtggagctggccacggagttcccagccacagaggcctgg 454
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db    301 GCAGGCTCCAGCACAGATGATGTGGAGCTGGCCACGGAGTTCCCAGCCACAGAGGCCTGG 360

Qy    455 gagtgtgagctagaaggcctgctggaagagaggcctgccctgtgcctgtccccgcaggcc 514
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GAGTGTGAGCTAGAAGGCCTGCTGGAAGAGAGGCCTGCCCTGTGCCTGTCCCCGCAGGCC 420

Qy    515 ccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcccagatctacatg 574
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CCATTTCCCAAGCTGGGCTGGGATGACGAACTGCGGAAACCCGGCGCCCAGATCTACATG 480

Qy    575 cgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagctagtcatc 634
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CGCTTCATGCAGGAGCACACCTGCTACGATGCCATGGCAACTAGCTCCAAGCTAGTCATC 540

Qy    635 ttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacggtgtgcgg 694
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TTCGACACCATGCTGGAGATCAAGAAGGCCTTCTTTGCTCTGGTGGCCAACGGTGTGCGG 600

Qy    695 gcagcccctctatgggacagcaagaagcagagctttgtggggatgctgaccatcactgac 754
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCAGCCCCTCTATGGGACAGCAAGAAGCAGAGCTTTGTGGGGATGCTGACCATCACTGAC 660

Qy    755 ttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatgagattgaa 814
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TTCATCCTGGTGCTGCATCGCTACTACAGGTCCCCCCTGGTCCAGATCTATGAGATTGAA 720

Qy    815 caacataagattgagacctggagggagatctacctgcaaggctgcttcaagcctctggtc 874
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CAACATAAGATTGAGACCTGGAGGGAGATCTACCTGCAAGGCTGCTTCAAGCCTCTGGTC 780

Qy    875 tccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaagaaccggatc 934
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TCCATCTCTCCTAATGATAGCCTGTTTGAAGCTGTCTACACCCTCATCAAGAACCGGATC 840

Qy    935 catcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctcacacacaaa 994
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CATCGCCTGCCTGTTCTTGACCCGGTGTCAGGCAACGTACTCCACATCCTCACACACAAA 900

Qy    995 cgcctgctcaagttcctgcacatctttggttccctgctgccccggcccctccttcctctac 1054
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CGCCTGCTCAAGTTCCTGCACATCTTTGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTAC 960

Qy   1055 cgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgctggagaca 1114
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CGCACTATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCTGTGGTGCTGGAGACA 1020

Qy   1115 gcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtg 1174
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GCACCCATCCTGACTGCACTGGACATCTTTGTGGACCGGCGTGTGTCTGCACTGCCTGTG 1080

Qy   1175 gtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattcacctggct 1234
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GTCAACGAATGTGGTCAGGTCGTGGGCCTCTATTCCCGCTTTGATGTGATTCACCTGGCT 1140

Qy   1235 gcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggcagaggaca 1294
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GCCCAGCAAACCTACAACCACCTGGACATGAGTGTGGGAGAAGCCCTGAGGCAGAGGACA 1200

Qy   1295 ctatgtctggagggagtcctttcctgccagccccacgagagcttgggggaagtgatcgac 1354
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTATGTCTGGAGGGAGTCCTTTCCTGCCAGCCCCACGAGAGCTTGGGGGAAGTGATCGAC 1260

Qy   1355 aggattgctcgggagcaggtacacaggctggtgctagtggacgagacccagcatctcttg 1414
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AGGATTGCTCGGGAGCAGGTACACAGGCTGGTGCTAGTGGACGAGACCCAGCATCTCTTG 1320

Qy   1415 ggcgtggtctccctctccgacatccttcaggcactggtgctcagccctgctggcatcgat 1474
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GGCGTGGTCTCCCTCTCCGACATCCTTCAGGCACTGGTGCTCAGCCCTGCTGGCATCGAT 1380

Qy   1475 gccctcggggcctgagaagatctgagtcctcaatcccaagccacctgcacacctggaagc 1534
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GCCCTCGGGGCCTGAGAAGATCTGAGTCCTCAATCCCAAGCCAACTGCACA-CTGGAAGC 1439
```

```
Qy   1535 caatgaagggaactggagaactcagccttcatcttcccccacccccatttgctggttcag 1594
          ||||||||| |    ||       |    |    |||||| ||||| ||||||||||||||
Db   1440 CAATGAAGGAATTGAGA------ACAGCTTCATTTCCCCAACCCCAATTTGCTGGTTCAG 1493

Qy   1595 ctatgattcaggtaggctctgccctgggccatgacaccagcct 1637
          ||||||||||||     || ||| |     ||  |    ||||
Db   1494 CTATGATTCAGGCTTCTTCAGCCTTCCAAAATTGCCTTTGCCT 1536


RESULT  4
AF214519
LOCUS       AF214519                2115 bp    mRNA    linear   PRI 03-JUN-2000
DEFINITION  Homo sapiens AMP-activated protein kinase gamma subunit (PRKAG3)
            mRNA, complete cds.
ACCESSION   AF214519
VERSION     AF214519.1  GI:8215681
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2115)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A., Thelander,M.,
            Rogel-Gaillard,C., Paul,S., Iannuccelli,N., Rask,L., Ronne,H.,
            Lundstrom,K., Reinsch,N., Gellin,J., Kalm,E., Roy,P.L., Chardon,P.
            and Andersson,L.
  TITLE     A mutation in PRKAG3 associated with excess glycogen content in pig
            skeletal muscle
  JOURNAL   Science 288 (5469), 1248-1251 (2000)
  MEDLINE   20280150
   PUBMED   10818001
REFERENCE   2  (bases 1 to 2115)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A.,
            Rogel-Gaillard,C., Paul,S., Gellin,J., Lundstrom,K., Reinsch,N.,
            Kalm,E., Le Roy,P., Chardon,P. and Andersson,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-1999) Dept Animal Breeding and Genetics, Swedish
            University of Agricultural Sciences, BMC box 597, Uppsala 751 24,
            Sweden
FEATURES             Location/Qualifiers
     source          1. .2115
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2p"
                     /tissue_type="skeletal muscle"
     gene            1. .2115
                     /gene="PRKAG3"
     CDS             1. .1395
                     /gene="PRKAG3"
                     /note="AMPKG3"
                     /codon_start=1
                     /product="AMP-activated protein kinase gamma subunit"
                     /protein_id="AAF73987.1"
                     /db_xref="GI:8215682"
                     /translation="MSFLEQENSSSWPSPAVTSSSERIRGKRRAKALRWTRQKSVEEG
                     EPPGQGEGPRSRPTAESTGLEATFPKTTPLAQADPAGVGTPPTGWDCLPSDCTASAAG
                     SSTDDVELATEFPATEAWECELEGLLEERPALCLSPQAPFPKLGWDDELRKPGAQIYM
                     RFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEQHKIETWREIYLQGCFKPLVSISPNDSLFEAVYTLI
                     KNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNECGQVVGLYSRFDVIHLAAQQTYNHLDMSV
                     GEALRQRTLCLEGVLSCQPHESLGEVIDRIARAEQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT      460 a    622 c    562 g    471 t
ORIGIN


  Query Match           88.2%;  Score 1453;  DB 9;  Length 2115;
  Best Local Similarity  97.3%;  Pred. No. 0;
  Matches 1501;  Conservative    0;  Mismatches   35;  Indels    7;  Gaps    2;
```

```
Qy     95 atgagcttcctagagcaagaaaacagcagctcatggccatcaccagctgtgaccagcagc 154
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGAGCTTCCTAGAGCAAGAAAACAGCAGCTCATGGCCATCACCAGCTGTGACCAGCAGC 60

Qy    155 tcagaaagaatccgtgggaaacggagggccaaagccttgagatggacaaggcagaagtcg 214
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 TCAGAAAGAATCCGTGGGAAACGGAGGGCCAAAGCCTTGAGATGGACAAGGCAGAAGTCG 120

Qy    215 gtggaggaaggggagccaccaggtcaggggaaggtccccggtccaggccagctgctgag 274
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db    121 GTGGAGGAAGGGGAGCCACCAGGTCAGGGGGAAGGTCCCCGGTCCAGGCCAACTGCTGAG 180

Qy    275 tccaccgggctggaggccacattccccaagaccacaccctggctcaagctgatcctgcc 334
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 TCCACCGGGCTGGAGGCCACATTCCCCAAGACCACACCCTTGGCTCAAGCTGATCCTGCC 240

Qy    335 ggggtgggcactccaccaacagggtgggactgcctcccctctgactgtacagcctcagct 394
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGGGTGGGCACTCCACCAACAGGGTGGGACTGCCTCCCCTCTGACTGTACAGCCTCAGCT 300

Qy    395 gcaggctccagcacagatgatgtggagctggccacggagttcccagccacagaggcctgg 454
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GCAGGCTCCAGCACAGATGATGTGGAGCTGGCCACGGAGTTCCCAGCCACAGAGGCCTGG 360

Qy    455 gagtgtgagctagaaggcctgctggaagagaggcctgccctgtgcctgtccccgcaggcc 514
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GAGTGTGAGCTAGAAGGCCTGCTGGAAGAGAGGCCTGCCCTGTGCCTGTCCCCGCAGGCC 420

Qy    515 ccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcccagatctacatg 574
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CCATTTCCCAAGCTGGGCTGGGATGACGAACTGCGGAAACCCGGCGCCCAGATCTACATG 480

Qy    575 cgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagctagtcatc 634
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CGCTTCATGCAGGAGCACACCTGCTACGATGCCATGGCAACTAGCTCCAAGCTAGTCATC 540

Qy    635 ttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacggtgtgcgg 694
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 TTCGACACCATGCTGGAGATCAAGAAGGCCTTCTTTGCTCTGGTGGCCAACGGTGTGCGG 600

Qy    695 gcagcccctctatgggacagcaagaagcagagctttgtggggatgctgaccatcactgac 754
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 GCAGCCCCTCTATGGGACAGCAAGAAGCAGAGCTTTGTGGGGATGCTGACCATCACTGAC 660

Qy    755 ttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatgagattgaa 814
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 TTCATCCTGGTGCTGCATCGCTACTACAGGTCCCCCCTGGTCCAGATCTATGAGATTGAA 720

Qy    815 caacataagattgagacctggagggagatctacctgcaaggctgcttcaagcctctggtc 874
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CAACATAAGATTGAGACCTGGAGGGAGATCTACCTGCAAGGCTGCTTCAAGCCTCTGGTC 780

Qy    875 tccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaagaaccggatc 934
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TCCATCTCTCCTAATGATAGCCTGTTTGAAGCTGTCTACACCCTCATCAAGAACCGGATC 840

Qy    935 catcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctcacacacaaa 994
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CATCGCCTGCCTGTTCTTGACCCGGTGTCAGGCAACGTACTCCACATCCTCACACACAAA 900

Qy    995 cgcctgctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctac 1054
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CGCCTGCTCAAGTTCCTGCACATCTTTGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTAC 960

Qy   1055 cgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgctggagaca 1114
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 CGCACTATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCTGTGGTGCTGGAGACA 1020

Qy   1115 gcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtg 1174
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GCACCCATCCTGACTGCACTGGACATCTTTGTGGACCGGCGTGTGTCTGCACTGCCTGTG 1080

Qy   1175 gtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattcacctggct 1234
```

```
Db   1081 GTCAACGAATGTGGTCAGGTCGTGGGCCTCTATTCCCGCTTTGATGTGATTCACCTGGCT 1140

Qy   1235 gcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggcagaggaca 1294
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GCCCAGCAAACCTACAACCACCTGGACATGAGTGTGGGAGAAGCCCTGAGGCAGAGGACA 1200

Qy   1295 ctatgtctggagggagtcctttcctgccagccccacgagagcttgggggaagtgatcgac 1354
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTATGTCTGGAGGGAGTCCTTTCCTGCCAGCCCCACGAGAGCTTGGGGGAAGTGATCGAC 1260

Qy   1355 aggattgctcgggagcaggtacacaggctggtgctagtggacgagacccagcatctcttg 1414
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AGGATTGCTCGGGAGCAGGTACACAGGCTGGTGCTAGTGGACGAGACCCAGCATCTCTTG 1320

Qy   1415 ggcgtggtctccctctccgacatccttcaggcactggtgctcagccctgctggcatcgat 1474
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GGCGTGGTCTCCCTCTCCGACATCCTTCAGGCACTGGTGCTCAGCCCTGCTGGCATCGAT 1380

Qy   1475 gccctcggggcctgagaagatctgagtcctcaatcccaagccacctgcacacctggaagc 1534
          |||||||||||||||||||||||||||||||||||||||| ||||||| ||||||||
Db   1381 GCCCTCGGGGCCTGAGAAGATCTGAGTCCTCAATCCCAAGCCAACTGCACA-CTGGAAGC 1439

Qy   1535 caatgaagggaactggagaactcagccttcatcttcccccaccccccatttgctggttcag 1594
          ||||||||| |    ||      |    |    |||||| ||||| ||||||||||||||
Db   1440 CAATGAAGGAATTGAGA------ACAGCTTCATTTCCCCAACCCCAATTTGCTGGTTCAG 1493

Qy   1595 ctatgattcaggtaggctctgccctgggccatgacaccagcct 1637
          ||||||||||||     || ||| |     ||  |    ||||
Db   1494 CTATGATTCAGGCTTCTTCAGCCTTCCAAAATTGCCTTTGCCT 1536


RESULT   5
AX099776
LOCUS       AX099776               2109 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 3 from Patent WO0120003.
ACCESSION   AX099776
VERSION     AX099776.1  GI:13538810
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2109)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 3 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .2109
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             472. .1389
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35799.1"
                     /db_xref="GI:13538811"
                     /translation="MRFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAP
                     LWDSKKQSFVGMLTITDFILVLHRYYRSPLVQIYEIEQHKIETWREIYLQGCFKPLVS
                     ISPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLPRPSFL
                     YRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSALPVVNECGQVVGLYSRFDVI
                     HLAAQQTYNHLDMSVGEALRQRTLCLEGVLSCQPHESLGEVIDRIAREQVHRLVLVDE
                     TQHLLGVVSLSDILQALVLSPAGIDALGA"
BASE COUNT      458 a    621 c    560 g    470 t
ORIGIN


  Query Match             87.9%;  Score 1447;  DB 6;  Length 2109;
  Best Local Similarity   97.3%;  Pred. No. 0;
```

```
  Matches 1495;  Conservative    0;  Mismatches   35;  Indels    7;  Gaps    2;

Qy    101 ttcctagagcaagaaaacagcagctcatggccatcaccagctgtgaccagcagctcagaa 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 TTCCTAGAGCAAGAAAACAGCAGCTCATGGCCATCACCAGCTGTGACCAGCAGCTCAGAA 60

Qy    161 agaatccgtgggaaacggagggccaaagccttgagatggacaaggcagaagtcggtggag 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 AGAATCCGTGGGAAACGGAGGGCCAAAGCCTTGAGATGGACAAGGCAGAAGTCGGTGGAG 120

Qy    221 gaaggggagccaccaggtcaggggggaaggtccccggtccaggccagctgctgagtccacc 280
          |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db    121 GAAGGGGAGCCACCAGGTCAGGGGGAAGGTCCCCGGTCCAGGCCAACTGCTGAGTCCACC 180

Qy    281 gggctggaggccacattccccaagaccacacccttggctcaagctgatcctgccggggtg 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 GGGCTGGAGGCCACATTCCCCAAGACCACACCCTTGGCTCAAGCTGATCCTGCCGGGGTG 240

Qy    341 ggcactccaccaacagggtgggactgcctcccctctgactgtacagcctcagctgcaggc 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGCACTCCACCAACAGGGTGGGACTGCCTCCCCTCTGACTGTACAGCCTCAGCTGCAGGC 300

Qy    401 tccagcacagatgatgtggagctggccacggagttcccagccacagaggcctgggagtgt 460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 TCCAGCACAGATGATGTGGAGCTGGCCACGGAGTTCCCAGCCACAGAGGCCTGGGAGTGT 360

Qy    461 gagctagaaggcctgctggaagagaggcctgccctgtgcctgtccccgcaggccccattt 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 GAGCTAGAAGGCCTGCTGGAAGAGAGGCCTGCCCTGTGCCTGTCCCCGCAGGCCCCATTT 420

Qy    521 cccaagctgggctgggatgacgaactgcggaaacccggcgcccagatctacatgcgcttc 580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CCCAAGCTGGGCTGGGATGACGAACTGCGGAAACCCGGCGCCCAGATCTACATGCGCTTC 480

Qy    581 atgcaggagcacacctgctacgatgccatggcaactagctccaagctagtcatcttcgac 640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 ATGCAGGAGCACACCTGCTACGATGCCATGGCAACTAGCTCCAAGCTAGTCATCTTCGAC 540

Qy    641 accatgctggagatcaagaaggccttctttgctctggtggccaacggtgtgcgggcagcc 700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ACCATGCTGGAGATCAAGAAGGCCTTCTTTGCTCTGGTGGCCAACGGTGTGCGGGCAGCC 600

Qy    701 cctctatgggacagcaagaagcagagctttgtggggatgctgaccatcactgacttcatc 760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CCTCTATGGGACAGCAAGAAGCAGAGCTTTGTGGGGATGCTGACCATCACTGACTTCATC 660

Qy    761 ctggtgctgcatcgctactacaggtccccccctggtccagatctatgagattgaacaacat 820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 CTGGTGCTGCATCGCTACTACAGGTCCCCCCTGGTCCAGATCTATGAGATTGAACAACAT 720

Qy    821 aagattgagacctggagggagatctacctgcaaggctgcttcaagcctctggtctccatc 880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 AAGATTGAGACCTGGAGGGAGATCTACCTGCAAGGCTGCTTCAAGCCTCTGGTCTCCATC 780

Qy    881 tctcctaatgatagcctgtttgaagctgtctacaccctcatcaagaaccggatccatcgc 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TCTCCTAATGATAGCCTGTTTGAAGCTGTCTACACCCTCATCAAGAACCGGATCCATCGC 840

Qy    941 ctgcctgttcttgacccggtgtcaggcaacgtactccacatcctcacacacaaacgcctg 1000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 CTGCCTGTTCTTGACCCGGTGTCAGGCAACGTACTCCACATCCTCACACACAAACGCCTG 900

Qy   1001 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 1060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 CTCAAGTTCCTGCACATCTTTGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACT 960

Qy   1061 atccaagatttgggcatcggcacattccgagacttggctgtggtgctggagacagcaccc 1120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 ATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCTGTGGTGCTGGAGACAGCACCC 1020

Qy   1121 atcctgactgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaac 1180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ATCCTGACTGCACTGGACATCTTTGTGGACCGGCGTGTGTCTGCACTGCCTGTGGTCAAC 1080
```

```
Qy   1181 gaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattcacctggctgcccag 1240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 GAATGTGGTCAGGTCGTGGGCCTCTATTCCCGCTTTGATGTGATTCACCTGGCTGCCCAG 1140

Qy   1241 caaacctacaaccacctggacatgagtgtgggagaagccctgaggcagaggacactatgt 1300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 CAAACCTACAACCACCTGGACATGAGTGTGGGAGAAGCCCTGAGGCAGAGGACACTATGT 1200

Qy   1301 ctggagggagtcctttcctgccagccccacgagagcttgggggaagtgatcgacaggatt 1360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGAGGGAGTCCTTTCCTGCCAGCCCCACGAGAGCTTGGGGGAAGTGATCGACAGGATT 1260

Qy   1361 gctcgggagcaggtacacaggctggtgctagtggacgagacccagcatctcttgggcgtg 1420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GCTCGGGAGCAGGTACACAGGCTGGTGCTAGTGGACGAGACCCAGCATCTCTTGGGCGTG 1320

Qy   1421 gtctccctctccgacatccttcaggcactggtgctcagccctgctggcatcgatgccctc 1480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GTCTCCCTCTCCGACATCCTTCAGGCACTGGTGCTCAGCCCTGCTGGCATCGATGCCCTC 1380

Qy   1481 ggggcctgagaagatctgagtcctcaatcccaagccacctgcacacctggaagccaatga 1540
          |||||||||||||||||||||||||||||||||| ||||||| ||||||||||||||||
Db   1381 GGGGCCTGAGAAGATCTGAGTCCTCAATCCCAAGCCAACTGCACA-CTGGAAGCCAATGA 1439

Qy   1541 agggaactggagaactcagccttcatcttccccacccccatttgctggttcagctatga 1600
          ||| |    ||      |    |    |||||| ||||| ||||||||||||||||||||
Db   1440 AGGAATTGAGA------ACAGCTTCATTTCCCCAACCCCAATTTGCTGGTTCAGCTATGA 1493

Qy   1601 ttcaggtaggctctgccctgggccatgacaccagcct 1637
          ||||||     || ||| |     ||  |    ||||
Db   1494 TTCAGGCTTCTTCAGCCTTCCAAAATTGCCTTTGCCT 1530


RESULT   6
AX099804
LOCUS       AX099804               2022 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 31 from Patent WO0120003.
ACCESSION   AX099804
VERSION     AX099804.1  GI:13538838
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 2022)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 31 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .2022
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
BASE COUNT      412 a    623 c    593 g    394 t
ORIGIN


  Query Match           71.2%;  Score 1172.8;  DB 6;  Length 2022;
  Best Local Similarity  85.7%;  Pred. No. 1.3e-257;
  Matches 1351;  Conservative    0;  Mismatches  217;  Indels    8;  Gaps    4;

Qy     53 acccctttcctggagcagccttgggggttctgagcatcaagagatgagcttcctagagcaa 112
          || || ||||||||||||||| |||||  || ||||||||||||||||||||||||||||
Db    109 ACTCCCTCCTGGAGCAGCCTCGGGGGACCTAAGCATCAAGAGATGAGCTTCCTAGAGCAA 168

Qy    113 gaaaacagcagctcatggccatcaccagctgtgaccagcagctcagaaagaatccgtggg 172
          | | | ||| | |||||||||||| | |||||| |||| ||||||||||||||| || ||||
Db    169 GGAGAGAGCCGTTCATGGCCATCCCGAGCTGTAACCACCAGCTCAGAAAGAAGCCATGGG 228

Qy    173 aaacggagggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagcca 232
          | | | ||  ||| ||||  ||||||||||||||| ||   || ||||||||| |||
Db    229 GACCAGGGGAACAAGGCCTCTAGATGGACAAGGCAGGAGGATGTAGAGGAAGGGGGGCCT 288

Qy    233 ccaggtcaggggggaaggtccccggtccaggccagctgctgagtccaccgggctggaggcc 292
          || || | | |||||||||||| |||||||||| |||||||||||||||||| |||||||
Db    289 CCGGGCCCGAGGGAAGGTCCCCAGTCCAGGCCAGTTGCTGAGTCCACCGGGCAGGAGGCC 348

Qy    293 acattccccaagaccacaccccttggctcaagctgatcc---tgccggggtgggcactcca 349
          |||||||||||| |||||||||||| ||||| | |||    |||| ||||| ||  ||
Db    349 ACATTCCCCAAGGCCACACCCTTGGCCCAAGCCGCTCCCTTGGCCGAGGTGGACAACCCC 408

Qy    350 ccaacagggtgggactgcctcccctctgactgtacagcctcagctgcaggctccagcaca 409
          ||||||| | ||||| |||||||||||||||| |||||||||  | | ||||| ||||
Db    409 CCAACAGAGCGGGACATCCTCCCCTCTGACTGTGCAGCCTCAGCCTCCGACTCCAACACA 468

Qy    410 gatgatgtggagctggccacggagttcccagccacagaggcctgggagtgtgagctagaa 469
          ||  || |||| |||| ||  |||||| ||||| | | ||| | || |  ||||||
Db    469 GACCATCTGGATCTGGGCATAGAGTTCTCAGCCTCGGCGGCGTCGGGGGATGAGCT---T 525

Qy    470 ggcctgctggaagagaggcctgccctgtgcctgtccccgcaggccccatttcccaagctg 529
          || ||| ||||||||| ||| |||| |||||  |||||  |||  |  || |||| ||||
Db    526 GGGCTGGTGGAAGAGAAGCCAGCCCCGTGCCCATCCCCAGAGGTGCTGTTACCCAGGCTG 585

Qy    530 ggctgggatgacgaactgcggaaacccggcgcccagatctacatgcgcttcatgcaggag 589
          |||||||||| || |||| ||| || || |||||| ||||||||| ||||||||||||||
Db    586 GGCTGGGATGATGAGCTGCAGAAGCCGGGGGCCCAGGTCTACATGCACTTCATGCAGGAG 645

Qy    590 cacacctgctacgatgccatggcaactagctccaagctagtcatcttcgacaccatgctg 649
          |||||||||||||||||||||||| || |||||||| || ||||||||||||||||||||||
Db    646 CACACCTGCTACGATGCCATGGCGACCAGCTCCAAACTGGTCATCTTCGACACCATGCTG 705

Qy    650 gagatcaagaaggccttctttgctctggtggccaacggtgtgcgggcagcccctctatgg 709
          |||||||||||||||||||||| |||||||||||||| || || || || ||| | |||
Db    706 GAGATCAAGAAGGCCTTCTTTGCCCTGGTGGCCAACGGCGTCCGAGCGGCACCTTTGTGG 765

Qy    710 gacagcaagaagcagagctttgtggggatgctgaccatcactgacttcatcctggtgctg 769
          |||||||||||||||||||| |||||||||||||||||||| |||||||| ||||||||
Db    766 GACAGCAAGAAGCAGAGCTTCGTGGGGATGCTGACCATCACAGACTTCATCTTGGTGCTG 825

Qy    770 catcgctactacaggtccccccctggtccagatctatgagattgaacaacataagattgag 829
          || ||||| |||||||||||||||||||||||||||||||| |||||||| ||||||||||
Db    826 CACCGCTATTACAGGTCCCCCCTGGTCCAGATCTACGAGATTGAAGAACATAAGATTGAG 885

Qy    830 acctggagggagatctacctgcaaggctgcttcaagcctctggtctccatctctcctaat 889
          |||||||||||||||||||| |||||||||||||||||||||||||||||||||| |||
Db    886 ACCTGGAGGGAGATCTACCTTCAAGGCTGCTTCAAGCCTCTGGTCTCCATCTCTCCCAAT 945

Qy    890 gatagcctgtttgaagctgtctacaccctcatcaagaaccggatccatcgcctgcctgtt 949
          || |||||||| |||||||||||| |||||||||||||||||||||| |||||||| ||
Db    946 GACAGCCTGTTCGAAGCTGTCTACGCCCTCATCAAGAACCGGATCCACCGCCTGCCGGTC 1005

Qy    950 cttgacccggtgtcaggcaacgtactccacatcctcacacacaaacgcctgctcaagttc 1009
          || ||||| || || ||    || ||||||||||||||||| || || || ||||||||||
Db   1006 CTGGACCCTGTCTCCGGGGCTGTGCTCCACATCCTCACACATAAGCGGCTTCTCAAGTTC 1065

Qy   1010 ctgcacatctttggttccctgctgccccggccctccttcctctaccgcactatccaagat 1069
          |||||||||||||  |||||||||||||||||||||||||||||||||| |||||||||
Db   1066 CTGCACATCTTTGGCACCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACCATCCAAGAT 1125

Qy   1070 ttgggcatcggcacattccgagacttggctgtggtgctggagacagcacccatcctgact 1129
          |||||||||||||||||||||||||||| |||||||||| || || ||||||||||
Db   1126 TTGGGCATCGGCACATTCCGAGACTTGGCCGTGGTGCTGGAAACGGCGCCCATCCTGACC 1185

Qy   1130 gcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtggt 1189
          |||||||||||||| |||||||||||||||||||||||||||||||||||||| |||
Db   1186 GCACTGGACATCTTCGTGGACCGGCGTGTGTCTGCGCTGCCTGTGGTCAACGAAACTGGA 1245

Qy   1190 caggtcgtgggcctctattcccgctttgatgtgattcacctggctgcccagcaaacctac 1249
          ||||| |||||||||||| || |||||||||||||||| |||||||||||||| ||||| |||
Db   1246 CAGGTAGTGGGCCTCTACTCTCGCTTTGATGTGATCCACCTGGCTGCCCAACAAACATAC 1305
```

```
Qy   1250 aaccacctggacatgagtgtgggagaagccctgaggcagaggacactatgtctggaggga 1309
          |||||||||||||||| ||||||||||||||||||||| ||||||| |||||||| ||
Db   1306 AACCACCTGGACATGAATGTGGGAGAAGCCCTGAGGCAGCGGACACTGTGTCTGGAAGGC 1365

Qy   1310 gtcctttcctgccagccccacgagagcttgggggaagtgatcgacaggattgctcgggag 1369
          |||||||||||||||||||||||| |||||||||||| || ||| ||||||  ||||||
Db   1366 GTCCTTTCCTGCCAGCCCCACGAGACCTTGGGGGAAGTCATTGACCGGATTGTCCGGGAA 1425

Qy   1370 caggtacacaggctggtgctagtggacgagacccagcatctcttgggcgtggtctccctc 1429
          |||||  ||| | |||||||| ||||| |||||||||||  ||  ||||||||| ||||||
Db   1426 CAGGTGCACCGCCTGGTGCTCGTGGATGAGACCCAGCACCTTCTGGGCGTGGTGTCCCTC 1485

Qy   1430 tccgacatccttcaggcactggtgctcagccctgctggcatcgatgccctcggggcctga 1489
          ||  |||||||||||||| |||||||||||||||||||| ||  |||||||||||||||||
Db   1486 TCTGACATCCTTCAGGCTCTGGTGCTCAGCCCTGCTGGAATTGATGCCCTCGGGGCCTGA 1545

Qy   1490 gaagatctgagtcctcaatcccaagccacct-gcacacctggaagccaatgaagggaact 1548
          |||  |  ||  | |   || || ||||||| |||||||||||||||| |||||||| |
Db   1546 GAACCTTGGAACCTTTGCTCTCAGGCCACCTGGCACACCTGGAAGCCAGTGAAGGGAGC- 1604

Qy   1549 ggagaactcagccttcatcttcccccacccccatttgctggttcagctatgattcaggta 1608
          | | |||||||  |||  | ||| || ||||| ||||||||   ||| |  ||||||||
Db   1605 CGTGGACTCAGCTCTCACTTCCCCTCAGCCCCACTTGCTGGTCTGGCTCTTGTTCAGGTA 1664

Qy   1609 ggctctgccctgggcc 1624
          |||||  |||| |||||
Db   1665 GGCTCCGCCCGGGGCC 1680


RESULT   7
AF214520
LOCUS       AF214520               1873 bp    mRNA    linear   MAM 03-JUN-2000
DEFINITION  Sus scrofa AMP-activated protein kinase gamma subunit (PRKAG3)
            mRNA, complete cds.
ACCESSION   AF214520
VERSION     AF214520.1  GI:8215683
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 1873)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A., Thelander,M.,
            Rogel-Gaillard,C., Paul,S., Iannuccelli,N., Rask,L., Ronne,H.,
            Lundstrom,K., Reinsch,N., Gellin,J., Kalm,E., Roy,P.L., Chardon,P.
            and Andersson,L.
  TITLE     A mutation in PRKAG3 associated with excess glycogen content in pig
            skeletal muscle
  JOURNAL   Science 288 (5469), 1248-1251 (2000)
  MEDLINE   20280150
   PUBMED   10818001
REFERENCE   2  (bases 1 to 1873)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A.,
            Rogel-Gaillard,C., Paul,S., Gellin,J., Lundstrom,K., Reinsch,N.,
            Kalm,E., Le Roy,P., Chardon,P. and Andersson,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-1999) Dept Animal Breeding and Genetics, Swedish
            University of Agricultural Sciences, BMC box 597, Uppsala 751 24,
            Sweden
FEATURES             Location/Qualifiers
     source          1. .1873
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /chromosome="15"
                     /map="15q"
                     /tissue_type="skeletal muscle"
     gene            1. .1873
                     /gene="PRKAG3"
     CDS             1. .1395
                     /gene="PRKAG3"
                     /note="AMPKG3"
                     /codon_start=1
                     /product="AMP-activated protein kinase gamma subunit"
                     /protein_id="AAF73988.1"
                     /db_xref="GI:8215684"
                     /translation="MSFLEQGESRSWPSRAVTTSSERSHGDQGNKASRWTRQEDVEEG
                     GPPGPREGPQSRPVAESTGQEATFPKATPLAQAAPLAEVDNPPTERDILPSDCAASAS
                     DSNTDHLDLGIEFSASAASGDELGLVEEKPAPCPSPEVLLPRLGWDDELQKPGAQVYM
                     HFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEEHKIETWREIYLQGCFKPLVSISPNDSLFEAVYALI
                     KNRIHRLPVLDPVSGAVLHILTHKRLLKFLHIFGTLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNETGQVVGLYSRFDVIHLAAQQTYNHLDMNV
                     GEALRQRTLCLEGVLSCQPHETLGEVIDRIVREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT      382 a    580 c    535 g    376 t
ORIGIN


  Query Match           69.2%;  Score 1140.4;  DB 4;  Length 1873;
  Best Local Similarity  85.7%;  Pred. No. 3.3e-250;
  Matches 1315;  Conservative    0;  Mismatches  211;  Indels    8;  Gaps    4;

Qy      95 atgagcttcctagagcaagaaaacagcagctcatggccatcaccagctgtgaccagcagc 154
           |||||||||||||||||| | | ||| | |||||||||| | |||||| |||| ||||
Db       1 ATGAGCTTCCTAGAGCAAGGAGAGAGCCGTTCATGGCCATCCCGAGCTGTAACCACCAGC 60

Qy     155 tcagaaagaatccgtgggaaacggagggccaaagccttgagatggacaaggcagaagtcg 214
           |||||||||| || |||| | | | ||  ||| |||| ||||||||||||||||| ||
Db      61 TCAGAAAGAAGCCATGGGGACCAGGGGAACAAGGCCTCTAGATGGACAAGGCAGGAGGAT 120

Qy     215 gtggaggaaggggagccaccaggtcaggggggaaggtccccggtccaggccagctgctgag 274
           || |||||||||| ||| || || | | |||||||||||| |||||||||| |||||||
Db     121 GTAGAGGAAGGGGGGCCTCCGGGCCCGAGGGAAGGTCCCCAGTCCAGGCCAGTTGCTGAG 180

Qy     275 tccaccgggctggaggccacattccccaagaccacacccttggctcaagctgatcc---t 331
           |||||||||| |||||||||||||||||||| |||||||||||||| ||||| | |||
Db     181 TCCACCGGGCAGGAGGCCACATTCCCCAAGGCCACACCCTTGGCCCAAGCCGCTCCCTTG 240

Qy     332 gccggggtgggcactccaccaacagggtgggactgcctcccctctgactgtacagcctca 391
           |||| ||||| ||  || |||||||| | |||||  |||||||||||||||| ||||||||
Db     241 GCCGAGGTGGACAACCCCCCAACAGAGCGGGACATCCTCCCCTCTGACTGTGCAGCCTCA 300

Qy     392 gctgcaggctccagcacagatgatgtggagctggccacggagttcccagccacagaggcc 451
           ||  | | ||||| ||||||  || |||| |||| ||  |||||| ||||| | | |||
Db     301 GCCTCCGACTCCAACACAGACCATCTGGATCTGGGCATAGAGTTCTCAGCCTCGGCGGCG 360

Qy     452 tgggagtgtgagctagaaggcctgctggaagagaggcctgccctgtgcctgtccccgcag 511
           | || |  ||||||    || ||| |||||||| ||| |||| |||||  |||||  ||
Db     361 TCGGGGGATGAGCT---TGGGCTGGTGGAAGAGAAGCCAGCCCCGTGCCCATCCCCAGAG 417

Qy     512 gccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcccagatctac 571
           |  |  || |||| |||||||||||||| || |||| ||| || || |||||| |||||
Db     418 GTGCTGTTACCCAGGCTGGGCTGGGATGATGAGCTGCAGAAGCCGGGGGCCCAGGTCTAC 477

Qy     572 atgcgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagctagtc 631
           |||| |||||||||||||||||||||||||||||||||||| || |||||||| || |||
Db     478 ATGCACTTCATGCAGGAGCACACCTGCTACGATGCCATGGCGACCAGCTCCAAACTGGTC 537

Qy     632 atcttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacggtgtg 691
           |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db     538 ATCTTCGACACCATGCTGGAGATCAAGAAGGCCTTCTTTGCCCTGGTGGCCAACGGCGTC 597

Qy     692 cgggcagccccctctatgggacagcaagaagcagagctttgtggggatgctgaccatcact 751
           || || || ||| | |||||||||||||||||||||| ||||||||||||||||||||
Db     598 CGAGCGGCACCTTTGTGGGACAGCAAGAAGCAGAGCTTCGTGGGGATGCTGACCATCACA 657

Qy     752 gacttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatgagatt 811
           |||||||| |||||||||| ||||| |||||||||||||||||||||||||| ||||||
Db     658 GACTTCATCTTGGTGCTGCACCGCTATTACAGGTCCCCCCTGGTCCAGATCTACGAGATT 717

Qy     812 gaacaacataagattgagacctggagggagatctacctgcaaggctgcttcaagcctctg 871
           ||| ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     718 GAAGAACATAAGATTGAGACCTGGAGGGAGATCTACCTTCAAGGCTGCTTCAAGCCTCTG 777

Qy     872 gtctccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaagaaccgg 931
           |||||||||||||| ||||| |||||||| |||||||||||| ||||||||||||||||
```

```
Db     778 GTCTCCATCTCTCCCAATGACAGCCTGTTCGAAGCTGTCTACGCCCTCATCAAGAACCGG 837

Qy     932 atccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctcacacac 991
           ||||| |||||||| || || ||||| || || ||    || ||||||||||||||||
Db     838 ATCCACCGCCTGCCGGTCCTGGACCCTGTCTCCGGGGCTGTGCTCCACATCCTCACACAT 897

Qy     992 aaacgcctgctcaagttcctgcacatctttggttccctgctgccccggccctccttcctc 1051
           || || || ||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db     898 AAGCGGCTTCTCAAGTTCCTGCACATCTTTGGCACCCTGCTGCCCCGGCCCTCCTTCCTC 957

Qy    1052 taccgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgctggag 1111
           |||||||| |||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db     958 TACCGCACCATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCCGTGGTGCTGGAA 1017

Qy    1112 acagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcactgcct 1171
           || || |||||||||||| |||||||||||||| |||||||||||||||||||| ||||||
Db    1018 ACGGCGCCCATCCTGACCGCACTGGACATCTTCGTGGACCGGCGTGTGTCTGCGCTGCCT 1077

Qy    1172 gtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattcacctg 1231
           |||||||||||  ||| ||||| |||||||||| || |||||||||||||| ||||||
Db    1078 GTGGTCAACGAAACTGGACAGGTAGTGGGCCTCTACTCTCGCTTTGATGTGATCCACCTG 1137

Qy    1232 gctgcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggcagagg 1291
           |||||||| ||||| |||||||||||||||||| ||||||||||||||||||||||| ||
Db    1138 GCTGCCCAACAAACATACAACCACCTGGACATGAATGTGGGAGAAGCCCTGAGGCAGCGG 1197

Qy    1292 acactatgtctggagggagtcctttcctgccagccccacgagagcttgggggaagtgatc 1351
           ||||| |||||||| || |||||||||||||||||||||||||| |||||||||||| ||
Db    1198 ACACTGTGTCTGGAAGGCGTCCTTTCCTGCCAGCCCCACGAGACCTTGGGGGAAGTCATT 1257

Qy    1352 gacaggattgctcgggagcaggtacacaggctggtgctagtggacgagacccagcatctc 1411
           ||| |||||||  ||||| ||||| ||| | |||||||| ||||| |||||||||| ||
Db    1258 GACCGGATTGTCCGGGAACAGGTGCACCGCCTGGTGCTCGTGGATGAGACCCAGCACCTT 1317

Qy    1412 ttgggcgtggtctccctctccgacatccttcaggcactggtgctcagccctgctggcatc 1471
           |||||||||| |||||||| |||||||||||||| |||||||||||||||||||| ||
Db    1318 CTGGGCGTGGTGTCCCTCTCTGACATCCTTCAGGCTCTGGTGCTCAGCCCTGCTGGAATT 1377

Qy    1472 gatgccctcggggcctgagaagatctgagtcctcaatcccaagccacct-gcacacctgg 1530
           |||||||||||||||||||||  |  ||  | |  || || ||||||| ||||||||||
Db    1378 GATGCCCTCGGGGCCTGAGAACCTTGGAACCTTTGCTCTCAGGCCACCTGGCACACCTGG 1437

Qy    1531 aagccaatgaagggaactggagaactcagccttcatcttcccccaccccccatttgctggt 1590
           |||||| |||||||| |  | | ||||||| |||  | ||| || ||||| ||||||||
Db    1438 AAGCCAGTGAAGGGAGC-CGTGGACTCAGCTCTCACTTCCCCTCAGCCCCACTTGCTGGT 1496

Qy    1591 tcagctatgattcaggtaggctctgccctgggcc 1624
              ||| |  |||||||||||||| |||| |||||
Db    1497 CTGGCTCTTGTTCAGGTAGGCTCCGCCCGGGGCC 1530


RESULT   8
AX099800
LOCUS       AX099800               1873 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 27 from Patent WO0120003.
ACCESSION   AX099800
VERSION     AX099800.1  GI:13538834
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 1873)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 27 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .1873
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
     CDS             1. .1395
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35800.1"
                     /db_xref="GI:13538835"
                     /translation="MSFLEQGESRSWPSRAVTTSSERSHGDQGNKASRWTRQEDVEEG
                     GPPGPREGPQSRPVAESTGQEATFPKATPLAQAAPLAEVDNPPTERDILPSDCAASAS
                     DSNTDHLDLGIEFSASAASGDELGLVEEKPAPCPSPEVLLPRLGWDDELQKPGAQVYM
                     HFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEEHKIETWREIYLQGCFKPLVSISPNDSLFEAVYALI
                     KNRIHRLPVLDPVSGAVLHILTHKRLLKFLHIFGTLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNETGQVVGLYSRFDVIHLAAQQTYNHLDMNV
                     GEALRQRTLCLEGVLSCQPHETLGEVIDRIVREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT      382 a    580 c    535 g    376 t
ORIGIN


  Query Match             69.2%;  Score 1140.4;  DB 6;  Length 1873;
  Best Local Similarity   85.7%;  Pred. No. 3.3e-250;
  Matches 1315;  Conservative    0;  Mismatches  211;  Indels    8;  Gaps    4;

Qy      95 atgagcttcctagagcaagaaaacagcagctcatggccatcaccagctgtgaccagcagc 154
           |||||||||||||||||||| | | ||| | |||||||||||| | |||||| |||| ||||
Db       1 ATGAGCTTCCTAGAGCAAGGAGAGAGCCGTTCATGGCCATCCCGAGCTGTAACCACCAGC 60

Qy     155 tcagaaagaatccgtgggaaacggagggccaaagccttgagatggacaaggcagaagtcg 214
           |||||||||| || |||| | | | ||  ||| ||||  |||||||||||||| ||
Db      61 TCAGAAAGAAGCCATGGGGACCAGGGGAACAAGGCCTCTAGATGGACAAGGCAGGAGGAT 120

Qy     215 gtggaggaaggggagccaccaggtcagggggaaggtccccggtccaggccagctgctgag 274
           || |||||||||| ||| || || | | |||||||||||| |||||||||| |||||||
Db     121 GTAGAGGAAGGGGGGCCTCCGGGCCCGAGGGAAGGTCCCCAGTCCAGGCCAGTTGCTGAG 180

Qy     275 tccaccgggctggaggccacattccccaagaccacacccttggctcaagctgatcc---t 331
           |||||||||| |||||||||||||||||| |||||||||||| ||||| | |||
Db     181 TCCACCGGGCAGGAGGCCACATTCCCCAAGGCCACACCCTTGGCCCAAGCCGCTCCCTTG 240

Qy     332 gccggggtgggcactccaccaacagggtgggactgcctcccctctgactgtacagcctca 391
           |||| ||||| ||  || ||||||| | |||||  ||||||||||||||| ||||||||
Db     241 GCCGAGGTGGACAACCCCCCCAACAGAGCGGGACATCCTCCCCTCTGACTGTGCAGCCTCA 300

Qy     392 gctgcaggctccagcacagatgatgtggagctggccacggagttcccagccacagaggcc 451
           ||  | | ||||| ||||||  || |||| |||| ||  |||||| |||| | | |||
Db     301 GCCTCCGACTCCAACACAGACCATCTGGATCTGGGCATAGAGTTCTCAGCCTCGGCGGCG 360

Qy     452 tgggagtgtgagctagaaggcctgctggaagagaggcctgccctgtgcctgtccccgcag 511
           | || |  ||||||    || ||| ||||||||| ||| |||| |||||  |||||  ||
Db     361 TCGGGGGATGAGCT---TGGGCTGGTGGAAGAGAAGCCAGCCCCGTGCCCATCCCCAGAG 417

Qy     512 gccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcccagatctac 571
           |  |  || |||| ||||||||||||||| || |||| ||| || || |||||| |||||
Db     418 GTGCTGTTACCCAGGCTGGGCTGGGATGATGAGCTGCAGAAGCCGGGGGCCCAGGTCTAC 477

Qy     572 atgcgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagctagtc 631
           |||| ||||||||||||||||||||||||||||||||||||||||| || |||||||| ||| |
Db     478 ATGCACTTCATGCAGGAGCACACCTGCTACGATGCCATGGCGACCAGCTCCAAACTGGTC 537

Qy     632 atcttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacggtgtg 691
           |||||||||||||||||||||||||||||||||||||||||||||| |||||||||||| ||
Db     538 ATCTTCGACACCATGCTGGAGATCAAGAAGGCCTTCTTTGCCCTGGTGGCCAACGGCGTC 597

Qy     692 cgggcagcccctctatgggacagcaagaagcagagctttgtggggatgctgaccatcact 751
           || || || ||| | |||||||||||||||||||||||||| ||||||||||||||||||
Db     598 CGAGCGGCACCTTTGTGGGACAGCAAGAAGCAGAGCTTCGTGGGGATGCTGACCATCACA 657

Qy     752 gacttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatgagatt 811
           ||||||||| |||||||||| ||||| |||||||||||||||||||||||||| ||||||
Db     658 GACTTCATCTTGGTGCTGCACCGCTATTACAGGTCCCCCCTGGTCCAGATCTACGAGATT 717

Qy     812 gaacaacataagattgagacctggagggagatctacctgcaaggctgcttcaagcctctg 871
```

```
Db    718 GAAGAACATAAGATTGAGACCTGGAGGGAGATCTACCTTCAAGGCTGCTTCAAGCCTCTG 777

Qy    872 gtctccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaagaaccgg 931
Db    778 GTCTCCATCTCTCCCAATGACAGCCTGTTCGAAGCTGTCTACGCCCTCATCAAGAACCGG 837

Qy    932 atccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctcacacac 991
Db    838 ATCCACCGCCTGCCGGTCCTGGACCCTGTCTCCGGGGCTGTGCTCCACATCCTCACACAT 897

Qy    992 aaacgcctgctcaagttcctgcacatctttggttccctgctgccccggccctccttcctc 1051
Db    898 AAGCGGCTTCTCAAGTTCCTGCACATCTTTGGCACCCTGCTGCCCCGGCCCTCCTTCCTC 957

Qy   1052 taccgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgctggag 1111
Db    958 TACCGCACCATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCCGTGGTGCTGGAA 1017

Qy   1112 acagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcactgcct 1171
Db   1018 ACGGCGCCCATCCTGACCGCACTGGACATCTTCGTGGACCGGCGTGTGTCTGCGCTGCCT 1077

Qy   1172 gtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattcacctg 1231
Db   1078 GTGGTCAACGAAACTGGACAGGTAGTGGGCCTCTACTCTCGCTTTGATGTGATCCACCTG 1137

Qy   1232 gctgcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggcagagg 1291
Db   1138 GCTGCCCAACAAACATACAACCACCTGGACATGAATGTGGGAGAAGCCCTGAGGCAGCGG 1197

Qy   1292 acactatgtctggagggagtcctttcctgccagccccacgagagcttgggggaagtgatc 1351
Db   1198 ACACTGTGTCTGGAAGGCGTCCTTTCCTGCCAGCCCCACGAGACCTTGGGGGAAGTCATT 1257

Qy   1352 gacaggattgctcgggagcaggtacacaggctggtgctagtggacgagacccagcatctc 1411
Db   1258 GACCGGATTGTCCGGGAACAGGTGCACCGCCTGGTGCTCGTGGATGAGACCCAGCACCTT 1317

Qy   1412 ttgggcgtggtctccctctccgacatccttcaggcactggtgctcagccctgctggcatc 1471
Db   1318 CTGGGCGTGGTGTCCCTCTCTGACATCCTTCAGGCTCTGGTGCTCAGCCCTGCTGGAATT 1377

Qy   1472 gatgccctcggggcctgagaagatctgagtcctcaatcccaagccacct-gcacacctgg 1530
Db   1378 GATGCCCTCGGGGCCTGAGAACCTTGGAACCTTTGCTCTCAGGCCACCTGGCACACCTGG 1437

Qy   1531 aagccaatgaagggaactggagaactcagccttcatcttccccacccccatttgctggt 1590
Db   1438 AAGCCAGTGAAGGGAGC-CGTGGACTCAGCTCTCACTTCCCCTCAGCCCCACTTGCTGGT 1496

Qy   1591 tcagctatgattcaggtaggctctgccctgggcc 1624
Db   1497 CTGGCTCTTGTTCAGGTAGGCTCCGCCCGGGGCC 1530
```

```
RESULT   9
AX099774
LOCUS       AX099774               1867 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 1 from Patent WO0120003.
ACCESSION   AX099774
VERSION     AX099774.1  GI:13538808
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 1867)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 1 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .1867
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
     CDS             472. .1389
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35798.1"
                     /db_xref="GI:13538809"
                     /translation="MHFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAP
                     LWDSKKQSFVGMLTITDFILVLHRYYRSPLVQIYEIEEHKIETWREIYLQGCFKPLVS
                     ISPNDSLFEAVYALIKNRIHRLPVLDPVSGAVLHILTHKRLLKFLHIFGTLLPRPSFL
                     YRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSALPVVNETGQVVGLYSRFDVI
                     HLAAQQTYNHLDMNVGEALRQRTLCLEGVLSCQPHETLGEVIDRIVREQVHRLVLVDE
                     TQHLLGVVSLSDILQALVLSPAGIDALGA"
BASE COUNT      380 a    583 c    529 g    375 t
ORIGIN
```

Query Match 68.9%; Score 1134.4; DB 6; Length 1867;
Best Local Similarity 85.7%; Pred. No. 7.7e-249;
Matches 1309; Conservative 0; Mismatches 211; Indels 8; Gaps 4;

```
Qy    101 ttcctagagcaagaaaacagcagctcatggccatcaccagctgtgaccagcagctcagaa 160
Db      1 TTCCTAGAGCAAGGAGAGAGCCGTTCATGGCCATCCCGAGCTGTAACCACCAGCTCAGAA 60

Qy    161 agaatccgtgggaaacggagggccaaagccttgagatggacaaggcagaagtcggtggag 220
Db     61 AGAAGCCATGGGGACCAGGGGAACAAGGCCTCTAGATGGACAAGGCAGGAGGATGTAGAG 120

Qy    221 gaaggggagccaccaggtcaggggggaaggtccccggtccaggccagctgctgagtccacc 280
Db    121 GAAGGGGGGCCTCCGGGCCCGAGGGAAGGTCCCCAGTCCAGGCCAGTTGCTGAGTCCACC 180

Qy    281 gggctggaggccacattccccaagaccacacccttggctcaagctgatcc---tgccggg 337
Db    181 GGGCAGGAGGCCACATTCCCCAAGGCCACACCCTTGGCCCAAGCCGCTCCCTTGGCCGAG 240

Qy    338 gtgggcactccaccaacagggtgggactgcctcccctctgactgtacagcctcagctgca 397
Db    241 GTGGACAACCCCCCAACAGAGCGGGACATCCTCCCCTCTGACTGTGCAGCCTCAGCCTCC 300

Qy    398 ggctccagcacagatgatgtggagctggccacggagttcccagccacagaggcctgggag 457
Db    301 GACTCCAACACAGACCATCTGGATCTGGGCATAGAGTTCTCAGCCTCGGCGGCGTCGGGG 360

Qy    458 tgtgagctagaaggcctgctggaagagaggcctgccctgtgcctgtccccgcaggcccca 517
Db    361 GATGAGCT---TGGGCTGGTGGAAGAGAAGCCAGCCCCGTGCCCATCCCCAGAGGTGCTG 417

Qy    518 tttcccaagctgggctgggatgacgaactgcggaaacccggcgcccagatctacatgcgc 577
Db    418 TTACCCAGGCTGGGCTGGGATGATGAGCTGCAGAAGCCGGGGGCCCAGGTCTACATGCAC 477

Qy    578 ttcatgcaggagcacacctgctacgatgccatggcaactagctccaagctagtcatcttc 637
Db    478 TTCATGCAGGAGCACACCTGCTACGATGCCATGGCGACCAGCTCCAAACTGGTCATCTTC 537

Qy    638 gacaccatgctggagatcaagaaggccttctttgctctggtggccaacggtgtgcgggca 697
Db    538 GACACCATGCTGGAGATCAAGAAGGCCTTCTTTGCCCTGGTGGCCAACGGCGTCCGAGCG 597

Qy    698 gcccctctatgggacagcaagaagcagagctttgtggggatgctgaccatcactgacttc 757
Db    598 GCACCTTTGTGGGACAGCAAGAAGCAGAGCTTCGTGGGGATGCTGACCATCACAGACTTC 657

Qy    758 atcctggtgctgcatcgctactacaggtccccccctggtccagatctatgagattgaacaa 817
Db    658 ATCTTGGTGCTGCACCGCTATTACAGGTCCCCCCTGGTCCAGATCTACGAGATTGAAGAA 717
```

```
Qy    818 cataagattgagacctggagggagatctacctgcaaggctgcttcaagcctctggtctcc 877
          ||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db    718 CATAAGATTGAGACCTGGAGGGAGATCTACCTTCAAGGCTGCTTCAAGCCTCTGGTCTCC 777

Qy    878 atctctcctaatgatagcctgtttgaagctgtctacaccctcatcaagaaccggatccat 937
          |||||||| ||||| |||||||| |||||||||||| ||||||||||||||||||||||
Db    778 ATCTCTCCCAATGACAGCCTGTTCGAAGCTGTCTACGCCCTCATCAAGAACCGGATCCAC 837

Qy    938 cgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctcacacacaaacgc 997
          |||||||| || || ||||| || || ||    || ||||||||||||||||| || ||
Db    838 CGCCTGCCGGTCCTGGACCCTGTCTCCGGGGCTGTGCTCCACATCCTCACACATAAGCGG 897

Qy    998 ctgctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgc 1057
          || ||||||||||||||||||||  |||||||||||||||||||||||||||||||||||
Db    898 CTTCTCAAGTTCCTGCACATCTTTGGCACCCTGCTGCCCCGGCCCTCCTTCCTCTACCGC 957

Qy    1058 actatccaagatttgggcatcggcacattccgagacttggctgtggtgctggagacagca 1117
           || ||||||||||||||||||||||||||||||||||||||||| |||||||||| || ||
Db     958 ACCATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCCGTGGTGCTGGAAACGGCG 1017

Qy    1118 cccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtc 1177
           |||||||||||| |||||||||||||| |||||||||||||||||||| ||||||||||||
Db    1018 CCCATCCTGACCGCACTGGACATCTTCGTGGACCGGCGTGTGTCTGCGCTGCCTGTGGTC 1077

Qy    1178 aacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattcacctggctgcc 1237
           |||||| ||| ||||| |||||||||| || |||||||||||||| ||||||||||||
Db    1078 AACGAAACTGGACAGGTAGTGGGCCTCTACTCTCGCTTTGATGTGATCCACCTGGCTGCC 1137

Qy    1238 cagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggcagaggacacta 1297
           || ||||| |||||||||||||||||| ||||||||||||||||||||| |||||||
Db    1138 CAACAAACATACAACCACCTGGACATGAATGTGGGAGAAGCCCTGAGGCAGCGGACACTG 1197

Qy    1298 tgtctggagggagtcctttcctgccagccccacgagagcttgggggaagtgatcgacagg 1357
           |||||||| || |||||||||||||||||||||||| |||||||||||| || ||| ||
Db    1198 TGTCTGGAAGGCGTCCTTTCCTGCCAGCCCCACGAGACCTTGGGGGAAGTCATTGACCGG 1257

Qy    1358 attgctcgggagcaggtacacaggctggtgctagtggacgagacccagcatctcttgggc 1417
           ||||  ||||| ||||| ||| | |||||||| ||||| ||||||||||| ||  |||||
Db    1258 ATTGTCCGGGAACAGGTGCACCGCCTGGTGCTCGTGGATGAGACCCAGCACCTTCTGGGC 1317

Qy    1418 gtggtctcccctctccgacatccttcaggcactggtgctcagccctgctggcatcgatgcc 1477
           ||||| |||||||| |||||||||||||| |||||||||||||||||||| || ||||||
Db    1318 GTGGTGTCCCTCTCTGACATCCTTCAGGCTCTGGTGCTCAGCCCTGCTGGAATTGATGCC 1377

Qy    1478 ctcggggcctgagaagatctgagtcctcaatcccaagccacct-gcacacctggaagcca 1536
           |||||||||||||||  |  ||  | |   || || ||||||| ||||||||||||||||
Db    1378 CTCGGGGCCTGAGAACCTTGGAACCTTTGCTCTCAGGCCACCTGGCACACCTGGAAGCCA 1437

Qy    1537 atgaagggaactggagaactcagccttcatcttccccccaccccatttgctggttcagct 1596
           |||||||| |  | | |||||||  |||  | ||| || ||||| ||||||||   |||
Db    1438 GTGAAGGGAGC-CGTGGACTCAGCTCTCACTTCCCCTCAGCCCCACTTGCTGGTCTGGCT 1496

Qy    1597 atgattcaggtaggctctgccctgggcc 1624
           |  |||||||||||||| |||| |||||
Db    1497 CTTGTTCAGGTAGGCTCCGCCCGGGGCC 1524


RESULT  10
RNAMPKGAM
LOCUS       RNAMPKGAM               1328 bp    mRNA    linear   ROD 25-FEB-1998
DEFINITION  R.norvegicus mRNA for gamma subunit of AMP-activated protein
            kinase.
ACCESSION   X95578
VERSION     X95578.1  GI:1185270
KEYWORDS    AMP-activated protein kinase; gamma subunit.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1328)
  AUTHORS   Woods,A., Cheung,P.C., Smith,F.C., Davison,M.D., Scott,J.,
            Beri,R.K. and Carling,D.
  TITLE     Characterization of AMP-activated protein kinase beta and gamma
            subunits. Assembly of the heterotrimeric complex in vitro
  JOURNAL   J. Biol. Chem. 271 (17), 10282-10290 (1996)
  MEDLINE   96215327
REFERENCE   2  (bases 1 to 1328)
  AUTHORS   Carling,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-FEB-1996) D. Carling, MRC Clinical Sciences Centre,
            Department of Molecular Medicine, RPMS, Hammersmith Hospital,
            DuCane Road, London, W12 0NN, UK
FEATURES             Location/Qualifiers
     source          1. .1328
                     /organism="Rattus norvegicus"
                     /strain="Wistar"
                     /db_xref="taxon:10116"
     CDS             37. .1029
                     /function="binds to alpha and beta subunits of AMPK,
                     forming a heterotrimer"
                     /codon_start=1
                     /product="AMP-activated protein kinase gamma"
                     /protein_id="CAA64831.1"
                     /db_xref="GI:1185271"
                     /db_xref="SWISS-PROT:P80385"
                     /translation="MESVAAESAPAPENEHSQETPESNSSVYTTFMKSHRCYDLIPTS
                     SKLVVFDTSLQVKKAFFALVTNGVRAAPLWDSKKQSFVGMLTITDFINILHRYYKSAL
                     VQIYELEEHKIETWREVYLQDSFKPLVCISPNASLFDAVSSLIRNKIHRLPVIDPESG
                     NTLYILTHKRILKFLKLFITEFPKPEFMSKSLEELQIGTYANIAMVRTTTPVYVALGI
                     FVQHRVSALPVVDEKGRVVDIYSKFDVINLAAEKTYNNLDVSVTKALQHRSHYFEGVL
                     KCYLHETLEAIINRLVEAEVHRLVVVDEHDVVKGIVSLSDILQALVLTGGEKKP"
BASE COUNT      330 a    337 c    340 g    321 t
ORIGIN


  Query Match           25.2%;  Score 414.8;  DB 10;  Length 1328;
  Best Local Similarity 64.9%;  Pred. No. 2.3e-84;
  Matches  614;  Conservative    0;  Mismatches  332;  Indels    0;  Gaps    0;

Qy    512 gccccatttcccaagctgggctgggatgacgaactgcggaaaccccggcgcccagatctac 571
          || ||| ||| ||  |      | | ||    | | || |   |       | |||
Db     61 GCTCCAGCTCCGGAGAATGAACACTCTCAAGAGACCCCGGAATCGAACAGTAGTGTGTAC 120

Qy    572 atgcgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagctagtc 631
          |    ||||||| ||   ||   |||||| ||    ||  | || ||||||||||| ||
Db    121 ACCACCTTCATGAAGTCTCATCGCTGCTATGACCTGATCCCCACAAGCTCCAAGCTGGTG 180

Qy    632 atcttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacggtgtg 691
          | || || ||   |||| || | ||||| ||||||||||| |||||| | ||||||||
Db    181 GTATTTGATACTTCGCTGCAGGTAAAGAAAGCCTTCTTTGCCCTGGTGACTAACGGTGTT 240

Qy    692 cgggcagcccctctatgggacagcaagaagcagagctttgtggggatgctgaccatcact 751
          || || |||||| | ||||| || ||||||||||||||||||||| ||||||||||||||
Db    241 CGTGCTGCCCCTTTGTGGGATAGTAAGAAGCAGAGCTTTGTGGGCATGCTGACCATCACT 300

Qy    752 gacttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatgagatt 811
          |||||||||    | ||||| || ||||||| |||  ||||||| ||||||||||  |
Db    301 GACTTCATCAATATTCTGCACCGATACTACAAGTCAGCCCTGGTGCAGATCTATGAACTG 360

Qy    812 gaacaacatagagattgagacctggagggagatctacctgcaaggctgcttcaagcctctg 871
          ||  | || ||||| ||||| ||||| ||| |||||||||||| || ||| ||||| ||
Db    361 GAGGAGCACAAGATAGAGACTTGGAGAGAGGTCTACCTGCAAGACTCCTTTAAGCCACTT 420

Qy    872 gtctccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaagaaccgg 931
          |||| ||| ||||| |||| ||| |||| || |||||||   |  | ||    ||   |
Db    421 GTCTGCATTTCTCCAAATGCCAGCTTGTTCGATGCTGTCTCTTCATTAATTCGAAATAAG 480

Qy    932 atccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctcacacac 991
          |||||  | || || ||| ||||||||| ||||||||||     |  |||| || || |||
Db    481 ATCCACAGGCTTCCAGTTATTGACCCGGAGTCAGGCAACACCTTGTACATTCTTACTCAC 540

Qy    992 aaacgcctgctcaagttcctgcacatctttggttccctgctgccccggccctccttcctc 1051
          || ||  | ||||||||||  |  | |||    |   | | ||| |||   ||| |
Db    541 AAGCGGATCCTCAAGTTCCTCAAGTTGTTTATCACTGAGTTCCCCAAGCCGGAATTCATG 600
```

```
Qy   1052 taccgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgctggag 1111
          |      || |  ||||  |    || ||||| | |    |  | ||  |||| |
Db    601 TCTAAGTCTCTGGAAGAGCTACAGATTGGCACCTACGCCAATATTGCCATGGTCCGTACC 660

Qy   1112 acagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcactgcct 1171
          ||  ||||  ||       || |||| |||||||||  | |  || || || ||  |||||
Db    661 ACTACACCTGTCTATGTGGCTCTGGGCATCTTTGTACAGCACCGAGTCTCCGCCTTGCCT 720

Qy   1172 gtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattcacctg 1231
          |||||  | ||    || |  || |||| | |||| |||   ||||||||||||| |  ||
Db    721 GTGGTGGATGAGAAAGGGCGTGTGGTGGACATCTACTCCAAGTTTGATGTGATTAATTTG 780

Qy   1232 gctgcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggcagagg 1291
          || ||  |  | || |||||| |||| ||  ||  ||||  | |||||||  |||  ||
Db    781 GCAGCAGAAAAGACATACAACAACCTAGATGTGTCTGTGACAAAAGCCCTACAGCACCGG 840

Qy   1292 acactatgtctggagggagtcctttcctgccagccccacgagagcttgggggaagtgatc 1351
           |||  |   | |||||  || ||    ||| | |  || ||||   | |  | | | |||
Db    841 TCACACTACTTCGAGGGTGTTCTCAAGTGCTACCTACATGAGACTCTAGAAGCAATCATC 900

Qy   1352 gacaggattgctcgggagcaggtacacaggctggtgctagtggacgagacccagcatctc 1411
           | ||  | |      |  |||| ||| | |||||| | ||||| ||      |     ||
Db    901 AATAGACTGGTGGAAGCAGAGGTTCACCGTCTGGTGGTGGTGGATGAACATGACGTGGTC 960

Qy   1412 ttgggcgtggtctccctctccgacatccttcaggcactggtgctca 1457
            |||| | || || || || |||||| | ||||| ||||||||||
Db    961 AAGGGCATTGTATCGCTGTCTGACATCTTACAGGCTCTGGTGCTCA 1006


RESULT  11
RNU42413
LOCUS       RNU42413                1550 bp    mRNA    linear   ROD 30-MAY-1996
DEFINITION  Rattus norvegicus 5'-AMP-activated protein kinase, gamma-1 subunit
            mRNA, complete cds.
ACCESSION   U42413
VERSION     U42413.1  GI:1335859
KEYWORDS    .
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 1550)
  AUTHORS   Gao,G., Fernandez,C.S., Stapleton,D., Auster,A.S., Widmer,J.,
            Dyck,J.R., Kemp,B.E. and Witters,L.A.
  TITLE     Non-catalytic beta- and gamma-subunit isoforms of the
            5'-AMP-activated protein kinase
  JOURNAL   J. Biol. Chem. 271 (15), 8675-8681 (1996)
  MEDLINE   96224074
REFERENCE   2  (bases 1 to 1550)
  AUTHORS   Gao,G., Widmer,J., Auster,A., Stapleton,D.S., Kemp,B.E. and
            Witters,L.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-DEC-1995) Lee A. Witters, Medicine/Biochemistry,
            Dartmouth Medical School, N. College St., Hanover, NH 03755-3833,
            USA
FEATURES             Location/Qualifiers
     source          1. .1550
                     /organism="Rattus norvegicus"
                     /db_xref="taxon:10116"
     CDS             <1. .972
                     /codon_start=1
                     /product="5'-AMP-activated protein kinase, gamma-1
                     subunit"
                     /protein_id="AAC52580.1"
                     /db_xref="GI:1335860"
                     /translation="SAPAPENEHSQETPESNSSVYTTFMKSHRCYDLIPTSSKLVVFD
                     TSLQVKKAFFALVTNGVRAAPLWDSKKQSFVGMLTITDFINILHRYYKSALVQIYELE
                     EHKIETWREVYLQDSFKPLVCISPNASLFDAVSSLIRNKIHRLPVIDPESGNTLYILT
                     HKRILKFLKLFITEFPKPEFMSKSLEELQIGTYANIAMVRTTTPVYVALGIFVQHRVS
                     ALPVVDEKGRVVDIYSKFDVINLAAEKTYNNLDVSVTKALQHRSHYFEGVLKCYLHET
                     LEAIINRLVEAEVHRLVVVDEHDVVKGIVSLSDILQALVLTGGEKKP"

BASE COUNT      371 a    398 c    394 g    387 t
ORIGIN

  Query Match           25.2%;  Score 414.8;  DB 10;  Length 1550;
  Best Local Similarity 64.9%;  Pred. No. 2.3e-84;
  Matches  614;  Conservative    0;  Mismatches  332;  Indels    0;  Gaps    0;

Qy     512 gccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcccagatctac 571
           || ||| ||| ||  |        | | ||    | | || |   |        | |||
Db       4 GCTCCAGCTCCGGAGAATGAACACTCTCAAGAGACCCCGGAATCGAACAGTAGTGTGTAC 63

Qy     572 atgcgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagctagtc 631
           |    ||||||| ||   ||  |||||| ||    ||  | || ||||||||||| ||
Db      64 ACCACCTTCATGAAGTCTCATCGCTGCTATGACCTGATCCCCACAAGCTCCAAGCTGGTG 123

Qy     632 atcttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacggtgtg 691
           | || || ||   |||| || | ||||| ||||||||||| |||||| | ||||||||
Db     124 GTATTTGATACTTCGCTGCAGGTAAAGAAAGCCTTCTTTGCCCTGGTGACTAACGGTGTT 183

Qy     692 cgggcagcccctctatgggacagcaagaagcagagctttgtggggatgctgaccatcact 751
           || || |||||| | ||||| || ||||||||||||||||||| ||||||||||||||||
Db     184 CGTGCTGCCCCTTTGTGGGATAGTAAGAAGCAGAGCTTTGTGGGCATGCTGACCATCACT 243

Qy     752 gacttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatgagatt 811
           |||||||||    | ||||| || ||||||| |||  ||||||| ||||||||||  |
Db     244 GACTTCATCAATATTCTGCACCGATACTACAAGTCAGCCCTGGTGCAGATCTATGAACTG 303

Qy     812 gaacaacataagattgagacctggagggagatctacctgcaaggctgcttcaagcctctg 871
           ||  | || ||||| ||||| ||||| ||| ||||||||||||| || ||| ||||| ||
Db     304 GAGGAGCACAAGATAGAGACTTGGAGAGAGGTCTACCTGCAAGACTCCTTTAAGCCACTT 363

Qy     872 gtctccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaagaaccgg 931
           |||| ||| |||| |||| ||| |||| || |||||||   |  | ||    ||  |
Db     364 GTCTGCATTTCTCCAAATGCCAGCTTGTTCGATGCTGTCTCTTCATTAATTCGAAATAAG 423

Qy     932 atccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctcacacac 991
           |||||  | || || ||| |||||||| ||||||||||    |  |||| || || |||
Db     424 ATCCACAGGCTTCCAGTTATTGACCCGGAGTCAGGCAACACCTTGTACATTCTTACTCAC 483

Qy     992 aaacgcctgctcaagttcctgcacatctttggttccctgctgccccggccctccttcctc 1051
           || ||  | ||||||||||  |  | |||    |  | | |||  |||    ||| |
Db     484 AAGCGGATCCTCAAGTTCCTCAAGTTGTTTATCACTGAGTTCCCCAAGCCGGAATTCATG 543

Qy    1052 taccgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgctggag 1111
           |      || |  ||||  |    || ||||| | |    |  | ||  |||| |
Db     544 TCTAAGTCTCTGGAAGAGCTACAGATTGGCACCTACGCCAATATTGCCATGGTCCGTACC 603

Qy    1112 acagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcactgcct 1171
           ||  ||||  ||       || |||| |||||||||  | |  || || || ||  |||||
Db     604 ACTACACCTGTCTATGTGGCTCTGGGCATCTTTGTACAGCACCGAGTCTCCGCCTTGCCT 663

Qy    1172 gtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattcacctg 1231
           |||||  | ||    || |  || |||| | |||| |||   ||||||||||||| |  ||
Db     664 GTGGTGGATGAGAAAGGGCGTGTGGTGGACATCTACTCCAAGTTTGATGTGATTAATTTG 723

Qy    1232 gctgcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggcagagg 1291
           || ||  |  | || |||||| |||| ||  ||  ||||  | |||||||  |||  ||
Db     724 GCAGCAGAAAAGACATACAACAACCTAGATGTGTCTGTGACAAAAGCCCTACAGCACCGG 783

Qy    1292 acactatgtctggagggagtcctttcctgccagccccacgagagcttgggggaagtgatc 1351
           |||  |   | |||||  || ||    ||| | |  || ||||   | |  | | | |||
Db     784 TCACACTACTTCGAGGGTGTTCTCAAGTGCTACCTACATGAGACTCTAGAAGCAATCATC 843

Qy    1352 gacaggattgctcgggagcaggtacacaggctggtgctagtggacgagacccagcatctc 1411
           | ||  | |      |  |||| ||| | |||||| | ||||| ||      |     ||
Db     844 AATAGACTGGTGGAAGCAGAGGTTCACCGTCTGGTGGTGGTGGATGAACATGACGTGGTC 903

Qy    1412 ttgggcgtggtctccctctccgacatccttcaggcactggtgctca 1457
             |||| | || || || || |||||| | ||||| ||||||||||
Db     904 AAGGGCATTGTATCGCTGTCTGACATCTTACAGGCTCTGGTGCTCA 949
```

```
RESULT 12
AX281579
LOCUS       AX281579                989 bp    DNA     linear   PAT 02-NOV-2001
DEFINITION  Sequence 2 from Patent WO0177305.
ACCESSION   AX281579
VERSION     AX281579.1  GI:16608830
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Andersson,L., Luthman,H. and Marklund,S.
  TITLE     Variants of the human amp-activated protein kinase gamma 3 subunit
  JOURNAL   Patent: WO 0177305-A 2 18-OCT-2001;
            Arexis AB (SE)
FEATURES             Location/Qualifiers
     source          1. .989
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      229 a    306 c    286 g    168 t
ORIGIN


  Query Match             24.7%;  Score 407.2;  DB 6;  Length 989;
  Best Local Similarity   99.3%;  Pred. No. 1.3e-82;
  Matches  409;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy     247 aggtccccggtccaggccagctgctgagtccaccgggctggaggccacattccccaagac 306
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     540 AGGTCCCCGGTCCAGGCCAGCTGCTGAGTCCACCGGGCTGGAGGCCACATTCCCCAAGAC 599

Qy     307 cacacccttggctcaagctgatcctgccggggtgggcactccaccaacagggtgggactg 366
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     600 CACACCCTTGGCTCAAGCTGATCCTGCCGGGGTGGGCACTCCACCAACAGGGTGGGACTG 659

Qy     367 cctcccctctgactgtacagcctcagctgcaggctccagcacagatgatgtggagctggc 426
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     660 CCTCCCCTCTGACTGTACAGCCTCAGCTGCAGGCTCCAGCACAGATGATGTGGAGCTGGC 719

Qy     427 cacggagttcccagccacagaggcctgggagtgtgagctagaaggcctgctggaagagag 486
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     720 CACGGAGTTCCCAGCCACAGAGGCCTGGGAGTGTGAGCTAGAAGGCCTGCTGGAAGAGAG 779

Qy     487 gcctgccctgtgcctgtccccgcaggccccatttcccaagctgggctgggatgacgaaact 546
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     780 GCCTGCCCTGTGCCTGTCCCCGCAGGCCCCATTTCCCAAGCTGGGCTGGGATGACGAACT 839

Qy     547 gcggaaacccggcgcccagatctacatgcgcttcatgcaggagcacacctgctacgatgc 606
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     840 GCGGAAACCCGGCGCCCAGATCTACATGCGCTTCATGCAGGAGCACACCTGCTACGATGC 899

Qy     607 catggcaactagctccaagctagtcatcttcgacaccatgctggagatcaag 658
           ||||||||||||||||||||||||||||||||||||||||||| | | |
Db     900 CATGGCAACTAGCTCCAAGCTAGTCATCTTCGACACCATGCTGGAGGTGAGG 951


RESULT 13
AC073128/c
LOCUS       AC073128              196554 bp    DNA     linear   HTG 21-FEB-2001
DEFINITION  Homo sapiens chromosome 2 clone RP11-647O5, WORKING DRAFT SEQUENCE,
            17 unordered pieces.
ACCESSION   AC073128
VERSION     AC073128.3  GI:13027579
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 196554)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 196554)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JUN-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Feb 21, 2001 this sequence version replaced gi:8469048.

            ---------------- Genome Center ----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: H_NH0647O05
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 98%
            Sequencing vector: plasmid; 0%
            Chemistry: Dye-primer ET; 98% of reads
            Chemistry: Dye-terminator Big Dye; 0% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 187795 bases at least Q40
            Consensus quality: 190513 bases at least Q30
            Consensus quality: 192099 bases at least Q20
            Insert size: 200000; agarose-fp
            Insert size: 194954; sum-of-contigs
            Quality coverage: 5.58 in Q20 bases; agarose-fp
            Quality coverage: 5.67 in Q20 bases; sum-of-contigs
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 17 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1157: contig of 1157 bp in length
            *     1158     1257: gap of unknown length
            *     1258     3600: contig of 2343 bp in length
            *     3601     3700: gap of unknown length
            *     3701     5103: contig of 1403 bp in length
            *     5104     5203: gap of unknown length
            *     5204     8524: contig of 3321 bp in length
            *     8525     8624: gap of unknown length
            *     8625    11856: contig of 3232 bp in length
            *    11857    11956: gap of unknown length
            *    11957    15783: contig of 3827 bp in length
            *    15784    15883: gap of unknown length
            *    15884    21906: contig of 6023 bp in length
            *    21907    22006: gap of unknown length
            *    22007    28887: contig of 6881 bp in length
            *    28888    28987: gap of unknown length
            *    28988    35255: contig of 6268 bp in length
            *    35256    35355: gap of unknown length
            *    35356    44642: contig of 9287 bp in length
            *    44643    44742: gap of unknown length
            *    44743    58275: contig of 13533 bp in length
            *    58276    58375: gap of unknown length
            *    58376    73816: contig of 15441 bp in length
            *    73817    73916: gap of unknown length
            *    73917    92140: contig of 18224 bp in length
            *    92141    92240: gap of unknown length
            *    92241   113337: contig of 21097 bp in length
            *   113338   113437: gap of unknown length
            *   113438   130325: contig of 16888 bp in length
            *   130326   130425: gap of unknown length
            *   130426   149287: contig of 18862 bp in length
            *   149288   149387: gap of unknown length
            *   149388   196554: contig of 47167 bp in length.
FEATURES             Location/Qualifiers
     source          1. .196554
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
```

```
                     /chromosome="2"
                     /clone="RP11-647O5"
     misc_feature    1. .1157
                     /note="assembly_name:Contig17"
     misc_feature    1258. .3600
                     /note="assembly_name:Contig18"
     misc_feature    3701. .5103
                     /note="assembly_name:Contig19
                     clone_end:T7
                     vector_side:right"
     misc_feature    5204. .8524
                     /note="assembly_name:Contig20"
     misc_feature    8625. .11856
                     /note="assembly_name:Contig21"
     misc_feature    11957. .15783
                     /note="assembly_name:Contig22"
     misc_feature    15884. .21906
                     /note="assembly_name:Contig23"
     misc_feature    22007. .28887
                     /note="assembly_name:Contig24"
     misc_feature    28988. .35255
                     /note="assembly_name:Contig25"
     misc_feature    35356. .44642
                     /note="assembly_name:Contig26"
     misc_feature    44743. .58275
                     /note="assembly_name:Contig27"
     misc_feature    58376. .73816
                     /note="assembly_name:Contig28"
     misc_feature    73917. .92140
                     /note="assembly_name:Contig29"
     misc_feature    92241. .113337
                     /note="assembly_name:Contig30"
     misc_feature    113438. .130325
                     /note="assembly_name:Contig31"
     misc_feature    130426. .149287
                     /note="assembly_name:Contig32"
     misc_feature    149388. .196554
                     /note="assembly_name:Contig33
                     clone_end:SP6
                     vector_side:right"
BASE COUNT    52296 a  46993 c  45889 g  49770 t   1606 others
ORIGIN


  Query Match             24.7%;  Score 407.2;  DB 2;  Length 196554;
  Best Local Similarity   99.3%;  Pred. No. 7.8e-83;
  Matches  409;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy     247 aggtccccggtccaggccagctgctgagtccaccgggctggaggccacattccccaagac 306
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62887 AGGTCCCCGGTCCAGGCCAGCTGCTGAGTCCACCGGGCTGGAGGCCACATTCCCCAAGAC 62828

Qy     307 cacacccttggctcaagctgatcctgccggggtgggcactccaccaacagggtgggactg 366
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62827 CACACCCTTGGCTCAAGCTGATCCTGCCGGGGTGGGCACTCCACCAACAGGGTGGGACTG 62768

Qy     367 cctcccctctgactgtacagcctcagctgcaggctccagcacagatgatgtggagctggc 426
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62767 CCTCCCCTCTGACTGTACAGCCTCAGCTGCAGGCTCCAGCACAGATGATGTGGAGCTGGC 62708

Qy     427 cacggagttcccagccacagaggcctgggagtgtgagctagaaggcctgctggaagagag 486
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62707 CACCGAGTTCCCAGCCACAGAGGCCTGGGAGTGTGAGCTAGAAGGCCTGCTGGAAGAGAG 62648

Qy     487 gcctgccctgtgcctgtccccgcaggccccatttcccaagctgggctgggatgacgaact 546
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62647 GCCTGCCCTGTGCCTGTCCCCGCAGGCCCCATTTCCCAAGCTGGGCTGGGATGACGAACT 62588

Qy     547 gcggaaacccggcgcccagatctacatgcgcttcatgcaggagcacacctgctacgatgc 606
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62587 GCGGAAACCCGGCGCCCAGATCTACATGCGCTTCATGCAGGAGCACACCTGCTACGATGC 62528

Qy     607 catggcaactagctccaagctagtcatcttcgacaccatgctggagatcaag 658
           ||||||||||||||||||||||||||||||||||||||||||||||||||| | | |
Db   62527 CATGGCAACTAGCTCCAAGCTAGTCATCTTCGACACCATGCTGGAGGTGAGG 62476


RESULT  14
AF036535
LOCUS       AF036535                1623 bp    mRNA    linear   ROD 30-MAY-2001
DEFINITION  Mus musculus AMP activated protein kinase mRNA, complete cds.
ACCESSION   AF036535
VERSION     AF036535.1  GI:2766684
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1623)
  AUTHORS   Shamsadin,R., Jantsan,K., Adham,I. and Engel,W.
  TITLE     Cloning, organisation, chromosomal localization and expression
            analysis of the mouse Prkag1 gene
  JOURNAL   Cytogenet. Cell Genet. 92 (1-2), 134-138 (2001)
  MEDLINE   21203559
   PUBMED   11306812
REFERENCE   2  (bases 1 to 1623)
  AUTHORS   Shamsadin,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-DEC-1997) AG.Engel, Humangenetik, Gosslerstr.12 d,
            Goettingen 37073, Germany
FEATURES             Location/Qualifiers
     source          1. .1623
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="15"
                     /tissue_type="testes"
     CDS             43. .1035
                     /note="AMPKg"
                     /codon_start=1
                     /product="AMP activated protein kinase"
                     /protein_id="AAB95475.1"
                     /db_xref="GI:2766685"
                     /translation="MESVAAESSPALENEHFQETPESNNSVYTSFMKSHRCYDLIPTS
                     SKLVVFDTSLQVKKAFFALVTNGVRAAPLWDSKKQCFVGMLTITDFINILHRYYKSAL
                     VQIYELEEHKIETWREVYLQDSFKPLVCISPNASSFDAVSSLIRNKIHRLPVIDPESG
                     NTLYILTHKRILKFLKLFIIEFPKPEFMSKSLQELQIGTYANIAMVRTTTPVYVALGI
                     FVQHRVSALPVVDEKGRVVDIYSKFDVINLAAEKTYNNLDVSVTKALXHRSHYFEGVL
                     KCYLHETLETIINRLVEAEVHRLVVVDEHXXVKGIVSLSDILQDLVLTGGEKKP"
BASE COUNT      383 a    414 c    416 g    400 t     10 others
ORIGIN


  Query Match             24.1%;  Score 396.8;  DB 10;  Length 1623;
  Best Local Similarity   65.2%;  Pred. No. 3e-80;
  Matches  581;  Conservative    0;  Mismatches  310;  Indels    0;  Gaps    0;

Qy     567 tctacatgcgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagc 626
           | || |    ||||||| ||   ||   |||||| ||    ||  | || || ||||||
Db     122 TGTATACTTCCTTCATGAAGTCTCATCGCTGCTATGACCTAATTCCCACAAGTTCCAAGT 181

Qy     627 tagtcatcttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacg 686
           | ||  | || |||||    |||  || | ||||| ||||| ||||| |||||| |||| |
Db     182 TGGTGGTATTTGACACTTCGCTACAGGTAAAGAAAGCCTTTTTTGCCCTGGTGACCAATG 241

Qy     687 gtgtgcgggcagcccctctatgggacagcaagaagcagagctttgtggggatgctgacca 746
           |||| || || |||||| | |||||||| ||||||||  | ||||||| ||||||||||
Db     242 GTGTTCGTGCCGCCCCTTTGTGGGACAGTAAGAAGCAGTGTTTTGTGGGCATGCTGACCA 301

Qy     747 tcactgacttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatg 806
           |||| |||||||||     |  |||| || ||||| | |||  ||||||| ||||| || |
Db     302 TCACCGACTTCATCAACATTTTGCACCGATACTATAAGTCAGCCCTGGTGCAGATTTACG 361

Qy     807 agattgaacaacataagattgagacctggagggagatctacctgcaaggctgcttcaagc 866
           |  | ||  | || ||||| ||||| ||||| ||| | |||||||| | || ||| ||||
Db     362 AACTCGAGGAGCACAAGATAGAGACGTGGAGAGAGGTGTACCTGCAGGACTCCTTTAAGC 421
```

```
Qy     867 ctctggtctccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaaga 926
Db     422 CACTTGTCTGCATCTCTCCAAATGCCAGCTCGTTTGATGCTGTCTCTTCATTAATTCGAA 481

Qy     927 accggatccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctca 986
Db     482 ATAAGATCCACAGGCTCCCAGTTATCGACCCAGAGTCAGGCAACACCTTGTACATCCTTA 541

Qy     987 cacacaaacgcctgctcaagttcctgcacatctttggttccctgctgccccggccctcct 1046
Db     542 CTCACAAGCGGATCCTCAAGTTCCTCAAGTTGTTTATCATCGAGTTCCCCAAGCCGGAAT 601

Qy    1047 tcctctaccgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgc 1106
Db     602 TCATGTCTAAGTCTCTCCAAGAGCTGCAGATTGGCACCTATGCCAATATTGCCATGGTCC 661

Qy    1107 tggagacagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcac 1166
Db     662 GTACTACCACGCCTGTCTACGTGGCTCTGGGCATCTTTGTACAGCACCGAGTCTCCGCCT 721

Qy    1167 tgcctgtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattc 1226
Db     722 TACCTGTAGTGGATGAGAAAGGGCGTGTGGTGGACATCTACTCCAAGTTTGATGTGATCA 781

Qy    1227 acctggctgcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggc 1286
Db     782 ATTTGGCAGCCGAAAAGACCTACAACAACCTAGATGTGTCTGTGACAAAAGCCCTGCANC 841

Qy    1287 agaggacactatgtctggagggagtcctttcctgccagccccacgagagcttgggggaag 1346
Db     842 ATCGGTCCCACTACTTTGAGGGTGTTCTCAAATGCTACCTGCATGAGACTCTGGAAACCA 901

Qy    1347 tgatcgacaggattgctcgggagcaggtacacaggctggtgctagtggacgagacccagc 1406
Db     902 TCATCAATAGGCTGGTGGAGGCAGAGGTTCACCGTCTGGTGGTGGTGGATGAACACNACN 961

Qy    1407 atctcttgggcgtggtctccctctccgacatccttcaggcactggtgctca 1457
Db     962 TGGTCAAGGGCATCGTTTCGCTGTCTGACATCTTACAGGATCTGGTGCTCA 1012
```

RESULT 15
BC015283

```
LOCUS       BC015283                3132 bp    mRNA    linear   ROD 04-OCT-2001
DEFINITION  Mus musculus, clone MGC:18882 IMAGE:4238045, mRNA, complete cds.
ACCESSION   BC015283
VERSION     BC015283.1  GI:15929719
KEYWORDS    MGC.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 3132)
  AUTHORS   Strausberg,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-OCT-2001) National Institutes of Health, Mammalian
            Gene Collection (MGC), Cancer Genomics Office, National Cancer
            Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
            USA
  REMARK    NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT     Contact: MGC help desk
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Sequencing Group at the Stanford Human Genome
            Center, Stanford University School of Medicine, Stanford, CA  94305
            Web site:       http://www-shgc.stanford.edu
            Contact:  (Dickson, Mark) mcd@paxil.stanford.edu
            Dickson, M., Schmutz, J., Grimwood, J., Rodriquez, A., and Myers,
            R. M.

            Clone distribution: MGC clone distribution information can be found
            through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
            Series: IRAK Plate: 25 Row: i Column: 18
            This clone was selected for full length sequencing because it
            passed the following selection criteria: Hexamer frequency ORF
            analysis, Similarity but not identity to protein.
FEATURES             Location/Qualifiers
     source          1. .3132
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="MGC:18882 IMAGE:4238045"
                     /tissue_type="Kidney, normal. 5 month old male mouse."
                     /clone_lib="NCI_CGAP_Kid14"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6"
     CDS             307. .2007
                     /codon_start=1
                     /product="Unknown (protein for MGC:18882)"
                     /protein_id="AAH15283.1"
                     /db_xref="GI:15929720"
                     /translation="MGSAAMDTKKKKEVSSPGGSSGKKNPSLKRRSLRVHIPDLSSFA
                     MPLLDGDVENSEKHSSRKVDSPFSSGSPSRGLFSRGPQPRPSSPVSAPVRPKTSPGSP
                     KTVFPFSYQESPPRSPRRMSFSGIFRSSSKESSPNSNPSTSPGGIRFFSRSRKTSSVS
                     SSPSTPTQVTKQHPFPLESYKQEPERPESRIYASSSPPDTGQRFCLAFQSPARPPLAS
                     PTYHAPLRTAVLAAAPGPAEAGMLEKLEFQEEEDSESGFYMRFMRSHKCYDIVPTSSK
                     LVVFDTTLQVKKAFFALVANGVRAAPLWESKKQSFVGMLTITDFINILHRYYKSPMVQ
                     IYELEEHKIETWRELYLQETFKPLVNISPDASLFDAVYSLIKNKIHRLPVIDPISGNA
                     LYILTHKRILKFLQLFMSDMPKPAFMKQNLDELGIGTYHNIAFIHPDTPIIKALNIFV
                     ERRISALPVVDESGKVVDIYSKFDVINLAAEKTYNNLDITVTQALQHRSQYFEGVVKC
                     SKLETLETIVDRIVRAEVHRLVVVNEADSIVGIISLSDILQALILTPAGAKQKETETE
                     "
BASE COUNT      822 a    802 c    755 g    753 t
ORIGIN
```

```
  Query Match             24.0%;  Score 395.4;  DB 10;  Length 3132;
  Best Local Similarity   64.9%;  Pred. No. 5.9e-80;
  Matches  585;  Conservative    0;  Mismatches  316;  Indels    0;  Gaps    0;

Qy     567 tctacatgcgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagc 626
Db    1076 TTTATATGCGATTCATGAGGTCACACAAGTGTTATGACATCGTTCCAACCAGTTCAAAGC 1135

Qy     627 tagtcatcttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacg 686
Db    1136 TTGTTGTCTTCGACACTACGTTGCAAGTCAAAAAGGCCTTCTTTGCCTTGGTAGCCAACG 1195

Qy     687 gtgtgcgggcagcccctctatgggacagcaagaagcagagctttgtggggatgctgacca 746
Db    1196 GAGTCCGTGCAGCGCCGCTGTGGGAAAGTAAAAAGCAGAGCTTTGTAGGAATGCTCACGA 1255

Qy     747 tcactgacttcatcctggtgctgcatcgctactacaggtccccctggtccagatctatg 806
Db    1256 TTACAGATTTCATAAATATCCTACACAGATACTATAAATCACCCATGGTACAGATTTATG 1315

Qy     807 agattgaacaacataagattgagacctggagggagatctacctgcaaggctgcttcaagc 866
Db    1316 AATTGGAAGAACATAAGATTGAAACGTGGAGGGAGTTGTACTTACAAGAAACCTTCAAGC 1375

Qy     867 ctctggtctccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaaga 926
Db    1376 CTTTGGTGAACATCTCTCCAGATGCGAGCCTCTTCGATGCTGTATACTCGTTGATCAAAA 1435

Qy     927 accggatccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctca 986
Db    1436 ATAAAATCCACAGATTGCCAGTTATTGACCCTATCAGTGGGAACGCACTTTATATACTTA 1495

Qy     987 cacacaaacgcctgctcaagttcctgcacatctttggttccctgctgccccggccctcct 1046
Db    1496 CCCACAAAAGAATCCTCAAGTTCCTCCAGCTTTTTATGTCTGACATGCCAAAGCCTGCCT 1555

Qy    1047 tcctctaccgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgc 1106
Db    1556 TCATGAAGCAGAACCTGGATGAGCTTGGAATCGGAACGTACCACAATATTGCCTTCATTC 1615
```

```
Qy   1107 tggagacagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcac 1166
               | |||||| | |   ||  || ||||||| |||||  | ||  | || |||
Db   1616 ACCCAGACACTCCCATCATCAAAGCCTTGAACATCTTCGTGGAGAGACGGATATCCGCAT 1675

Qy   1167 tgcctgtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattc 1226
          ||||||||||  | || |  ||  | || || |   | ||||||   |||||||| |||
Db   1676 TGCCTGTGGTGGATGAGTCAGGAAAAGTTGTAGATATTTATTCCAAGTTTGATGTAATTA 1735

Qy   1227 acctggctgcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggc 1286
          | || |||||  || |||| |||||  || |||||||  |  |||    | ||||||  ||
Db   1736 ATCTTGCTGCTGAGAAAACGTACAATAACTTGGACATCACAGTGACCCAGGCCCTGCAGC 1795

Qy   1287 agaggacactatgtctggagggagtcctttcctgccagccccacgagagcttgggggaag 1346
          |  |  | |  | | | ||||| ||  |    |||       |  || |   ||| |
Db   1796 ACCGCTCGCAGTATTTTGAGGGTGTGGTGAAGTGCAGTAAGCTGGAAACACTGGAGACCA 1855

Qy   1347 tgatcgacaggattgctcgggagcaggtacacaggctggtgctagtggacgagacccagc 1406
          |  | |||||||| |   | |   |||| ||  |||||||  ||||| | ||  |  |
Db   1856 TCGTGGACAGGATAGTGAGAGCCGAGGTCCATCGGCTGGTTGTAGTGAATGAAGCAGATA 1915

Qy   1407 atctcttgggcgtggtctccctctccgacatccttcaggcactggtgctcagccctgctg 1466
             |  ||||  |  ||||||| || |||||||| || || ||| | ||||  || || |
Db   1916 GCATTGTGGGTATTATCTCCCTGTCAGACATCCTGCAAGCCCTGATCCTCACACCAGCAG 1975

Qy   1467 g 1467
          |
Db   1976 G 1976
```

Search completed: October 3, 2002, 14:48:55
Job time: 12055 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 11:32:25 ; Search time 5701.1 Seconds
(without alignments)
4076.714 Million cell updates/sec

Title: US-09-826-581-3
Perfect score: 1722
Sequence: 1 cctggccccctcagatcaaga.....gatgagaggctcgggctgga 1722

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	148	8.6	572	10	BI344527	BI344527 373008 MA
	2	132.4	7.7	413	9	AA178898	AA178898 zp38d10.r
	3	116.8	6.8	1042	12	CNS03KJ0	AL248229 Tetraodon
	4	82.6	4.8	933	10	BG919314	BG919314 602817782
	5	82.2	4.8	633	10	BJ072114	BJ072114 BJ072114
	6	81.8	4.8	536	10	BM488662	BM488662 pgm2n.pk0
	7	80.2	4.7	595	10	BM487789	BM487789 pgm2n.pk0
	8	80.2	4.7	636	10	BG713637	BG713637 pg11n.pk0
	9	80.2	4.7	647	10	BM440762	BM440762 pgr1n.pk0
	10	80.2	4.7	649	9	AJ395115	AJ395115 AJ395115
	11	78.6	4.6	758	9	AJ396118	AJ396118 AJ396118
	12	76.4	4.4	576	9	AV603335	AV603335 AV603335
	13	73	4.2	564	9	AV608257	AV608257 AV608257
	14	71	4.1	775	10	BI833269	BI833269 603087149
c	15	70.4	4.1	448	9	AA558845	AA558845 nl69b09.s
	16	70.4	4.1	450	10	BF351397	BF351397 CM4-HT050
	17	70.4	4.1	455	9	AA578219	AA578219 nl56h03.s
	18	70.4	4.1	469	9	AL047390	AL047390 DKFZp586A
	19	70.4	4.1	473	9	BE166881	BE166881 CM4-HT050
	20	70.4	4.1	508	9	BE166874	BE166874 CM4-HT050
c	21	70.4	4.1	583	9	AW379936	AW379936 RC4-HT025
	22	70.4	4.1	586	10	BG609812	BG609812 323791 MA
	23	70.4	4.1	591	9	AW410926	AW410926 fh09c01.x
c	24	70.4	4.1	598	9	BE148626	BE148626 MR0-HT024
	25	70.4	4.1	616	9	AW956906	AW956906 EST368976
	26	70.4	4.1	634	10	BG740148	BG740148 602630747
	27	70.4	4.1	668	10	BG705895	BG705895 602669396
	28	70.4	4.1	708	9	AW411228	AW411228 fh11b02.x
	29	70.4	4.1	710	10	BI223706	BI223706 602943249
	30	70.4	4.1	726	10	BI768590	BI768590 603056916
	31	70.4	4.1	742	10	BI914634	BI914634 603179401
	32	70.4	4.1	746	10	BF528081	BF528081 602042828
	33	70.4	4.1	756	10	BI859947	BI859947 603386088
	34	70.4	4.1	782	10	BI819312	BI819312 603037761
	35	70.4	4.1	782	10	BI821538	BI821538 603038457
	36	70.4	4.1	795	10	BE871189	BE871189 601448767
	37	70.4	4.1	808	10	BI771184	BI771184 603054888
	38	70.4	4.1	826	9	AL519198	AL519198 AL519198
	39	70.4	4.1	841	10	BI546514	BI546514 603191524
	40	70.4	4.1	864	10	BI858240	BI858240 603384001
	41	70.4	4.1	870	10	BI910928	BI910928 603069326
	42	70.4	4.1	903	10	BG177822	BG177822 602314222
	43	70.4	4.1	905	9	AL555228	AL555228 AL555228
	44	70.4	4.1	908	9	AL552459	AL552459 AL552459
	45	70.4	4.1	921	9	AL548987	AL548987 AL548987

ALIGNMENTS

RESULT 1
BI344527
LOCUS BI344527 572 bp mRNA linear EST 30-JUL-2001
DEFINITION 373008 MARC 2PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BI344527
VERSION BI344527.1 GI:15037807
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 572)
AUTHORS Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W. and Keele,J.W.
TITLE Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR PRimers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCCAGTCACGACG
Plate: 119 row: I column: 11
Seq primer: ATTTAGGTGACACTATAG.
FEATURES Location/Qualifiers
source 1. .572
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2PIG"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;

Library made from pooled tissue from testis, ovary, endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 106 a 190 c 151 g 124 t 1 others
ORIGIN

Query Match 8.6%; Score 148; DB 10; Length 572;
Best Local Similarity 88.9%; Pred. No. 1.9e-22;
Matches 160; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

```
Qy   1510 ccatcctaaccagggttccctgctgccccggccctccttcctctaccgcactatccaaga 1569
          || | | |   ||  ||||||||||||||||||||||||||||||||| ||||||||
Db    171 CCTGCACATCTTTGGCACCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACCATCCAAGA 230

Qy   1570 tttgggcatcggcacattccgagacttggctgtggtgctggagacagcacccatcctgac 1629
          |||||||||||||||||||||||||||||| |||||||||| || || ||||||||||
Db    231 TTTGGGCATCGGCACATTCCGAGACTTGGCCGTGGTGCTGGAAACGGCGCCCATCCTGAC 290

Qy   1630 tgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtgg 1689
          |||||||||||||| |||||||||||||||||||| |||||||||||||||||  |||
Db    291 CGCACTGGACATCTTCGTGGACCGGCGTGTGTCTGCGCTGCCTGTGGTCAACGAAACTGG 350
```

RESULT 2
AA178898
LOCUS AA178898 413 bp mRNA linear EST 09-MAR-1998
DEFINITION zp38d10.r1 Stratagene muscle 937209 Homo sapiens cDNA clone IMAGE:611731 5' similar to SW:AAKG_RAT P80385 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA CHAIN ;, mRNA sequence.
ACCESSION AA178898
VERSION AA178898.1 GI:1760259
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 413)
AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
TITLE WashU-NCI human EST Project
JOURNAL Unpublished (1997)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1280 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 255.
FEATURES Location/Qualifiers
source 1. .413
/organism="Homo sapiens"
/db_xref="GDB:4643570"
/db_xref="taxon:9606"
/clone="IMAGE:611731"
/clone_lib="Stratagene muscle 937209"
/tissue_type="muscle"
/dev_stage="adult"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: skeletal muscle; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally. Primer: Oligo dT. Skeletal muscle from patient with malignant hyperthermia. Average insert size: 1.0 kb; Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG 3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTTT 3'"
BASE COUNT 80 a 109 c 117 g 102 t 5 others
ORIGIN

Query Match 7.7%; Score 132.4; DB 9; Length 413;
Best Local Similarity 97.3%; Pred. No. 4.8e-19;
Matches 144; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```
Qy   1543 ctccttcctctaccgcactatccaagatttgggcatcggcacattccgagacttggctgt 1602
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CTCCTTCCTCTACCGCACTATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCTGT 60

Qy   1603 ggtgctggagacagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtc 1662
          |||||||||||||   || ||||||||||||||||||||||||||||||||||||||||
Db     61 GGTGCTGGAGACA-NNCCAATCCTGACTGCACTGGACATCTTTGTGGACCGGCGTGTGTC 119

Qy   1663 tgcactgcctgtggtcaacgaatgtggt 1690
          |||||||||||||||||||||||||||
Db    120 TGCACTGCCTGTGGTCAACGAATGTGGT 147
```

RESULT 3
CNS03KJ0
LOCUS CNS03KJ0 1042 bp DNA linear GSS 17-MAY-2000
DEFINITION Tetraodon nigroviridis genome survey sequence T7 end of clone 033L13 of library G from Tetraodon nigroviridis, genomic survey sequence.
ACCESSION AL248229
VERSION AL248229.1 GI:7969241
KEYWORDS GSS; genome survey sequence.
SOURCE Tetraodon nigroviridis.
ORGANISM Tetraodon nigroviridis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.
REFERENCE 1 (bases 1 to 1042)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissenbach,J.
TITLE Charaterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1042)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissenbach,J.
TITLE Human gene number estimate provided by genome wide analysis using Tetraodon nigroviridis DNA sequence
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1042)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) to the EMBL/GenBank/DDBJ databases
COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon.
FEATURES Location/Qualifiers
source 1. .1042
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="033L13"
/clone_lib="G"
/note="Genoscope sequence ID : C0BG033CF07LP1-end : T7"
BASE COUNT 242 a 282 c 241 g 271 t 6 others
ORIGIN

Query Match 6.8%; Score 116.8; DB 12; Length 1042;
Best Local Similarity 56.3%; Pred. No. 2.1e-15;
Matches 269; Conservative 0; Mismatches 192; Indels 17; Gaps 2;

```
Qy   1238 tgcagcctgtttgaagctgtctacaccctcatcaagaaccggatccatcgcctgcctgtt 1297
          |||||||||||||| ||||| ||||| |||||||| |||   || || ||||||||||
Db    196 TGCAGCCTGTTTGATGCTGTGTACACGCTCATCAAAAACAAAATTCACCGCCTGCCTGTC 255
```

```
Qy   1298 cttgacccggtgtcaggcaacgtactccacatcctcacacacaaacgcctgctcaagttc 1357
          |||| || ||  |||| |||| |||  ||||||||||||||||  |  | |||||||||
Db    256 ATTGATCCAGTCACAGGAAACGCACTTTACATCCTCACACACAAGAGGATCCTCAAGTTC 315

Qy   1358 ctgcacatctttgtaag---------cctgggcccaggtgggaggaaggggagacctgg 1408
          || ||  | ||||| ||          |      |   ||| ||  |||| | |
Db    316 CTCCAGCTGTTTGTGAGTGCGACACATCCCCATCAGTGTGTGACAAAGGCGCCGTTGAAT 375

Qy   1409 gcaggtgatcagagggcctgaggagtcttcagccctagcagtcgtggggaagagctggga 1468
          |        | ||||       |       ||  |       ||  | |||   ||
Db    376 CCCACCCTGGTGGGGGCAGATCTTCGCCCATCTGCTTCCCCCTAGTGGCCAAAGCCCAGA 435

Qy   1469 gccctcttgaagctgctg--------gatccctgatctccacctggtccccatcctaacc 1520
          ||  ||    || || |         |  |  |      |  ||   | || |  ||
Db    436 CCCTCCTGTTTGCGGCCGTGAACACACCCCTGTCTTACTTTGCGCGTTTGCCTCGTTTCC 495

Qy   1521 agggttccctgctgccccggccctccttcctctaccgcactatccaagatttgggcatcg 1580
          ||   |    |  |||   |||| | ||| |  | |  ||  |    ||  ||||||| |
Db    496 AGATGTGTGAGAAGCCAAAGCCCGCTTTCATGAAGCAGACCCTGAGGGAGCTGGGCATTG 555

Qy   1581 gcacattccgagacttggctgtggtgctggagacagcacccatcctgactgcactggaca 1640
          | |||| ||| ||| | ||| |  | |         | ||||||  | |  || ||  |||
Db    556 GTACATACCGGGACATTGCTTTCATTCACCCCGACACGCCCATCATCAAAGCGCTCAACA 615

Qy   1641 tctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtggtacccaccc 1698
          ||||||||||  |||| ||||| || ||||| |||||  |||| |  ||||   ||||
Db    616 TCTTTGTGGAGAGGCGGGTGTCGGCCCTGCCGGTGGTGGACGACTCCGGTAGGAACCC 673
```

```
RESULT   4
BG919314
LOCUS       BG919314                933 bp    mRNA    linear   EST 05-JUN-2001
DEFINITION  602817782F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:4946661 5',
            mRNA sequence.
ACCESSION   BG919314
VERSION     BG919314.1  GI:14299790
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 933)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey Green M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10895  row: d  column: 22
            High quality sequence stop: 498.
FEATURES             Location/Qualifiers
     source          1. .933
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4946661"
                     /clone_lib="NCI_CGAP_Mam6"
                     /sex="female, virgin"
                     /tissue_type="infiltrating ductal carcinoma"
                     /dev_stage="5 months"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Jeffrey Green, M.D., NIH"
BASE COUNT      244 a    292 c    238 g    159 t
```

ORIGIN

```
  Query Match           4.8%;  Score 82.6;  DB 10;  Length 933;
  Best Local Similarity   52.1%;  Pred. No. 7.2e-08;
  Matches  232;  Conservative    0;  Mismatches  209;  Indels    4;  Gaps    2;

Qy   1242 gcctgtttgaagctgtctacaccctcatcaagaaccggatccatcgcctgcctgttcttg 1301
          | | | |||| |||||||   |  | ||    ||   ||||||  | || || ||| | |
Db      7 GTCCGGTTGATGCTGTCTCTTCATTAATTCGAAATAAGATCCACAGGCTCCCAGTTATCG 66

Qy   1302 acccggtgtcaggcaacgtactccacatcctcacacacaaacgcctgctcaagttcctgc 1361
          |||| | ||||||||||    |  |||||||| || ||||| ||  | ||||||||||
Db     67 ACCCAGAGTCAGGCAACACCTTGTACATCCTTACTCACAAGCGGATCCTCAAGTTCCTCA 126

Qy   1362 acatctttgtaag---cctgggcccaggtgggaggaaggggagacctgggcaggtgatc 1418
          |  | ||||||||   |||  | |||           | | |  || |||   |     |
Db    127 AGTTGTTTGTAAGTAACCTCAGGCCATCATCCGATTACGTGCTGAACTGCAGACACACAC 186

Qy   1419 agagggcctgaggagtcttcagccctagcagtcgtggggaagagctgggagccctcttga 1478
          |||| |  || |||      |||  |    ||  ||  ||     |   ||||
Db    187 AGAGCGGTTGGGGACAGAGGAGCGGTGTTGGTGTTGATGATTGAGCGTAGCACCTCCATG 246

Qy   1479 agctgctggatccctgatctccacctggtccccatcctaaccagggttccctgctgcccc 1538
          ||  |          | || ||       ||  | |   |||    ||  | | |||
Db    247 CTCT-ATCCGGAAAGCACCTGCAAGCCAATGCCCCCTTCCTCAGATCACCGAGTTCCCCA 305

Qy   1539 ggccctccttcctctaccgcactatccaagatttgggcatcggcacattccgagacttgg 1598
          |||    ||| | |      || |||||||  ||   || |||| |      |  | |
Db    306 AGCCGGAATTCATGTCTAAGTCTCTCCAAGAGCTGCAGATTGGCACCTATGCCAATATTG 365

Qy   1599 ctgtggtgctggagacagcacccatcctgactgcactggacatctttgtggaccggcgtg 1658
          |  |||| |     ||  | ||  ||      || |||| |||| |||| | |  || |
Db    366 CCATGGTCCGTACTACCACGCCTGTCTACGTGGCTCTGGGCATCCTTGTACAGCACCGAG 425

Qy   1659 tgtctgcactgcctgtggtcaacga 1683
          | || ||  | |||| ||  | ||
Db    426 TCTCCGCCTTACCTGTAGTGGATGA 450
```

```
RESULT   5
BJ072114
LOCUS       BJ072114                633 bp    mRNA    linear   EST 11-DEC-2001
DEFINITION  BJ072114 NIBB Mochii normalized Xenopus tailbud library Xenopus
            laevis cDNA clone XL096j16 5', mRNA sequence.
ACCESSION   BJ072114
VERSION     BJ072114.1  GI:17502303
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 633)
  AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
            ,Y.
  TITLE     Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .633
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="XL096j16"
                     /clone_lib="NIBB Mochii normalized Xenopus tailbud
                     library"
                     /tissue_type="whole embryo"
```

```
                     /dev_stage="stage 25"
BASE COUNT       144 a    162 c    148 g    179 t
ORIGIN


  Query Match             4.8%;  Score 82.2;  DB 10;  Length 633;
  Best Local Similarity   73.4%;  Pred. No. 7.4e-08;
  Matches  105;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Qy    1228 ctcccatctctgcagcctgtttgaagctgtctacaccctcatcaagaaccggatccatcg 1287
           ||||||       |||||| ||  ||||||| ||  | ||||||||||||   |||||| ||
Db      10 CTCCCCAGCAGACAGCCTTTTCCAAGCTGTGTATTCACTCATCAAGAATAAGATCCACCG 69

Qy    1288 cctgcctgttcttgacccggtgtcaggcaacgtactccacatcctcacacacaaacgcct 1347
           || || ||  | || || || || |||||| |  | || |||||||| || ||||||||
Db      70 GCTACCAGTAATGGATCCCGTATCTGGCAACATCTTGCATATCCTCACGCATAAACGCCT 129

Qy    1348 gctcaagttcctgcacatctttg 1370
           ||||||||||| ||| ||||||
Db     130 ACTCAAGTTCCTTCACCTCTTTG 152


RESULT   6
BM488662
LOCUS       BM488662                 536 bp    mRNA    linear   EST 07-FEB-2002
DEFINITION  pgm2n.pk008.g21 Normalized Chicken Breast Muscle, Leg Muscle, and
            Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cDNA
            clone pgm2n.pk008.g21 5' similar to gb|AAC52580.1 (U42413)
            5'-AMP-activated protein kinase, gamma-1 subunit [Rattus norvegicus
            ], mRNA sequence.
ACCESSION   BM488662
VERSION     BM488662.1  GI:18609593
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 536)
  AUTHORS   Cogburn,L.A. and Monsonego-Ornan,E.
  TITLE     ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and
            Epiphyseal Growth Plate cDNA library, USDA/IFAFS Animal Genome
            Project
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Larry A. Cogburn
            University of Delaware
            Townsend Hall, Newark, DE 19717, USA
            Tel: 302-831-1335
            Fax: 302-831-2822
            Email: cogburn@udel.edu, www.chickest.udel.edu.
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Gallus gallus"
                     /strain="Commercial broiler and  Ottawa Res. Centre
                     Strains 90 & 21"
                     /db_xref="taxon:9031"
                     /clone="pgm2n.pk008.g21"
                     /clone_lib="Normalized Chicken Breast Muscle, Leg Muscle,
                     and Epiphyseal Growth Plate cDNA library (pgm2n)"
                     /sex="Male and Female"
                     /tissue_type="Breast muscle, leg muscle and epiphyseal
                     growth plate"
                     /dev_stage="Breast,leg:Embryo(d19);post-hatch(1d,1,3,5,7,9
                     ,11 weeks);growth plate(1d,7d,14d post-hatch)"
                     /lab_host="E. coli EMDH10B"
                     /note="Vector: pCMVSPORT6; Library made from equivalent
                     pools of total RNA isolated from each tissue (embryonic
                     muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth
                     plate 33.3% of the final RNA pool). Single pass sequencing
                     from 5'-end"
BASE COUNT       117 a    171 c    132 g    116 t
ORIGIN


  Query Match             4.8%;  Score 81.8;  DB 10;  Length 536;
  Best Local Similarity   70.1%;  Pred. No. 8.4e-08;
  Matches  110;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy    1228 ctcccatctctgcagcctgtttgaagctgtctacaccctcatcaagaaccggatccatcg 1287
           |||||    |  |||||| ||||| || |||| | |||| |||   ||   |||||| ||
Db      85 CTCCCCCAACGCCAGCCTTTTTGATGCCGTCTCCTCCCTGATCCGCAATAAGATCCACCG 144

Qy    1288 cctgcctgttcttgacccggtgtcaggcaacgtactccacatcctcacacacaaacgcct 1347
           ||| || ||  | ||||| |  || ||||||   ||| |||||||||| |||||||| |
Db     145 CCTCCCCGTCATCGACCCCGACTCGGGCAACACTCTCTACATCCTCACCCACAAACGCAT 204

Qy    1348 gctcaagttcctgcacatctttgtaagcctgggccca 1384
           |||||||||||  |  ||||| |      || ||||
Db     205 CCTCAAGTTCCTCAAACTCTTTATTGCAGAGGTCCCA 241


RESULT   7
BM487789
LOCUS       BM487789                 595 bp    mRNA    linear   EST 07-FEB-2002
DEFINITION  pgm2n.pk005.j24 Normalized Chicken Breast Muscle, Leg Muscle, and
            Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cDNA
            clone pgm2n.pk005.j24 5' similar to gb|AAC52580.1 (U42413)
            5'-AMP-activated protein kinase, gamma-1 subunit [Rattus norvegicus
            ], mRNA sequence.
ACCESSION   BM487789
VERSION     BM487789.1  GI:18608720
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 595)
  AUTHORS   Cogburn,L.A. and Monsonego-Ornan,E.
  TITLE     ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and
            Epiphyseal Growth Plate cDNA library, USDA/IFAFS Animal Genome
            Project
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Larry A. Cogburn
            University of Delaware
            Townsend Hall, Newark, DE 19717, USA
            Tel: 302-831-1335
            Fax: 302-831-2822
            Email: cogburn@udel.edu, www.chickest.udel.edu.
FEATURES             Location/Qualifiers
     source          1. .595
                     /organism="Gallus gallus"
                     /strain="Commercial broiler and  Ottawa Res. Centre
                     Strains 90 & 21"
                     /db_xref="taxon:9031"
                     /clone="pgm2n.pk005.j24"
                     /clone_lib="Normalized Chicken Breast Muscle, Leg Muscle,
                     and Epiphyseal Growth Plate cDNA library (pgm2n)"
                     /sex="Male and Female"
                     /tissue_type="Breast muscle, leg muscle and epiphyseal
                     growth plate"
                     /dev_stage="Breast,leg:Embryo(d19);post-hatch(1d,1,3,5,7,9
                     ,11 weeks);growth plate(1d,7d,14d post-hatch)"
                     /lab_host="E. coli EMDH10B"
                     /note="Vector: pCMVSPORT6; Library made from equivalent
                     pools of total RNA isolated from each tissue (embryonic
                     muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth
                     plate 33.3% of the final RNA pool). Single pass sequencing
                     from 5'-end"
BASE COUNT       126 a    199 c    137 g    133 t
ORIGIN


  Query Match             4.7%;  Score 80.2;  DB 10;  Length 595;
  Best Local Similarity   69.4%;  Pred. No. 2e-07;
  Matches  109;  Conservative    0;  Mismatches   48;  Indels    0;  Gaps    0;
```

```
Qy   1228 ctcccatctctgcagcctgtttgaagctgtctacaccctcatcaagaaccggatccatcg 1287
          |||||       |||||| ||||| || |||| | |||| |||   ||   |||||| ||
Db    289 CTCCCCCAATGCCAGCCTTTTTGATGCCGTCTCCTCCCTGATCCGCAATAAGATCCACCG 348

Qy   1288 cctgcctgttcttgacccggtgtcaggcaacgtactccacatcctcacacacaaacgcct 1347
          ||| || ||  | ||||| |  || ||||||   ||| |||||||||| |||||||| |
Db    349 CCTCCCCGTCATCGACCCCGACTCGGGCAACACTCTCTACATCCTCACCCACAAACGCAT 408

Qy   1348 gctcaagttcctgcacatctttgtaagcctgggccca 1384
           ||||||||||  |  ||||| |      || ||||
Db    409 CCTCAAGTTCCTCAAACTCTTTATTGCAGAGGTCCCA 445


RESULT   8
BG713637
LOCUS       BG713637                636 bp    mRNA    linear   EST 08-MAY-2001
DEFINITION  pg11n.pk008.c13 Normalized Liver Library Gallus gallus cDNA clone
            pg11n.pk008.c13 5' similar to gi|4506061 ref|NP_002724.1| protein
            kinase, AMP-activated, gamma 1 non-catalytic subunit; AMPK gamma 1;
            Protein kinase, AMP-activated, noncatalytic, gemma-1 [Homo sapiens]
            gi|12737489 ref|XP_006778.2| protein kinase, AMP-activated, gamma 1
            non, mRNA sequence.
ACCESSION   BG713637
VERSION     BG713637.1  GI:14007587
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 636)
  AUTHORS   Burnside,J., Morgan,R.W. and Cogburn,L.A.
  TITLE     Chicken ESTs from a normalized liver library
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Joan Burnside
            Molecular Endocrinology
            University of Delaware
            40 Townsend Hall, Newark, DE 19717, USA
            Tel: 302 831-1345
            Fax: 302-831-3411
            Email: joan@UDel.Edu, www.chickest.udel.edu.
FEATURES             Location/Qualifiers
     source          1. .636
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /clone="pg11n.pk008.c13"
                     /clone_lib="Normalized Liver Library"
                     /sex="Male and Female"
                     /tissue_type="liver"
                     /lab_host="E.coli EMDH10B"
                     /note="Vector: pCMVSPORT 6"
BASE COUNT      129 a    215 c    167 g    119 t      6 others
ORIGIN


  Query Match             4.7%;  Score 80.2;  DB 10;  Length 636;
  Best Local Similarity   69.4%;  Pred. No. 2.1e-07;
  Matches  109;  Conservative    0;  Mismatches   48;  Indels    0;  Gaps    0;

Qy   1228 ctcccatctctgcagcctgtttgaagctgtctacaccctcatcaagaaccggatccatcg 1287
          |||||       |||||| ||||| || |||| | |||| |||   ||   |||||| ||
Db    429 CTCCCCCAATGCCAGCCTTTTTGATGCCGTCTCCTCCCTGATCCGCAATAAGATCCACCG 488

Qy   1288 cctgcctgttcttgacccggtgtcaggcaacgtactccacatcctcacacacaaacgcct 1347
          ||| || ||  | ||||| |  || ||||||   ||| |||||||||| |||||||| |
Db    489 CCTCCCCGTCATCGACCCCGACTCGGGCAACACTCTCTACATCCTCACCCACAAACGCAT 548

Qy   1348 gctcaagttcctgcacatctttgtaagcctgggccca 1384
           ||||||||||  |  ||||| |      || ||||
Db    549 CCTCAAGTTCCTCAAACTCTTTATTGCAGAGGTCCCA 585


RESULT   9
BM440762
LOCUS       BM440762                647 bp    mRNA    linear   EST 01-FEB-2002
DEFINITION  pgr1n.pk002.19 Normalized Chicken Reproductive Tract cDNA Library
            (pgr1n) Gallus gallus cDNA clone pgr1n.pk002.19 5' similar to
            gi|4506061 ref|NP_002724.1| protein kinase, AMP-activated, gamma 1
            non-catalytic subunit; AMPK gamma 1; Protein kinase, AMP-activated,
            noncatalytic, gemma-1 [Homo sapiens] gi|12737489 ref|XP_006778.2|
            protein kinase, AMP-activated, gamma 1, mRNA sequence.
ACCESSION   BM440762
VERSION     BM440762.1  GI:18471537
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 647)
  AUTHORS   Cogburn,L.A. and Nys,Y.
  TITLE     ESTs from Normalized Chicken Reproductive Tract cDNA library-
            University of Delaware and INRA, Tours-Poultry Unit Project
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Larry A. Cogburn
            University of Delaware
            Townsend Hall, Newark, DE 19717, USA
            Tel: 302-831-1335
            Fax: 302-831-2822
            Email: cogburn@udel.edu, www.chickest.udel.edu.
FEATURES             Location/Qualifiers
     source          1. .647
                     /organism="Gallus gallus"
                     /strain="Commercial broiler and layer"
                     /db_xref="taxon:9031"
                     /clone="pgr1n.pk002.19"
                     /clone_lib="Normalized Chicken Reproductive Tract cDNA
                     Library (pgr1n)"
                     /sex="Male and Female"
                     /tissue_type="Testis, ovary and oviduct"
                     /dev_stage="Various stages;embryonic, post-hatch, immature
                     and sexually-mature"
                     /lab_host="E. coli EMDH10B"
                     /note="Vector: pCMVSPORT6; Library made from three total
                     RNA pools from each tissue (testis 25%, ovary 25%, and
                     oviduct 50% of final RNA pool); Single pass sequencing
                     from 5'-end"
BASE COUNT      137 a    222 c    160 g    128 t
ORIGIN


  Query Match             4.7%;  Score 80.2;  DB 10;  Length 647;
  Best Local Similarity   69.4%;  Pred. No. 2.1e-07;
  Matches  109;  Conservative    0;  Mismatches   48;  Indels    0;  Gaps    0;

Qy   1228 ctcccatctctgcagcctgtttgaagctgtctacaccctcatcaagaaccggatccatcg 1287
          |||||       |||||| ||||| || |||| | |||| |||   ||   |||||| ||
Db    393 CTCCCCCAATGCCAGCCTTTTTGATGCCGTCTCCTCCCTGATCCGCAATAAGATCCACCG 452

Qy   1288 cctgcctgttcttgacccggtgtcaggcaacgtactccacatcctcacacacaaacgcct 1347
          ||| || ||  | ||||| |  || ||||||   ||| |||||||||| |||||||| |
Db    453 CCTCCCCGTCATCGACCCCGACTCGGGCAACACTCTCTACATCCTCACCCACAAACGCAT 512

Qy   1348 gctcaagttcctgcacatctttgtaagcctgggccca 1384
           ||||||||||  |  ||||| |      || ||||
Db    513 CCTCAAGTTCCTCAAACTCTTTATTGCAGAGGTCCCA 549


RESULT  10
AJ395115
LOCUS       AJ395115                649 bp    mRNA    linear   EST 25-JAN-2001
DEFINITION  AJ395115 dkfz426 Gallus gallus cDNA clone 21c2r1, mRNA sequence.
ACCESSION   AJ395115
VERSION     AJ395115.1  GI:7125706
KEYWORDS    EST.
```

```
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 649)
  AUTHORS   Abdrakhmanov,I., Lodygin,D., Geroth,P., Arakawa,H., Law,A., Plachy
            ,J., Korn,B. and Buerstedde,J.M.
  TITLE     A large database of chicken bursal ESTs as a resource for the
            analysis of vertebrate gene function
  JOURNAL   Genome Res. 10 (12), 2062-2069 (2000)
  MEDLINE   20568495
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinistr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES             Location/Qualifiers
     source          1. .649
                     /organism="Gallus gallus"
                     /strain="CB"
                     /db_xref="taxon:9031"
                     /clone="21c2r1"
                     /clone_lib="dkfz426"
                     /tissue_type="Bursa of Fabricius"
BASE COUNT      130 a    227 c    155 g    137 t
ORIGIN
```

```
  Query Match             4.7%;  Score 80.2;  DB 9;  Length 649;
  Best Local Similarity   69.4%;  Pred. No. 2.1e-07;
  Matches  109;  Conservative    0;  Mismatches   48;  Indels    0;  Gaps    0;

Qy   1228 ctcccatctctgcagcctgtttgaagctgtctacaccctcatcaagaaccggatccatcg 1287
          |||||       |||||| ||||| || |||| | |||| |||   ||   |||||| ||
Db     46 CTCCCCCAATGCCAGCCTTTTTGATGCCGTCTCCTCCCTGATCCGCAATAAGATCCACCG 105

Qy   1288 cctgcctgttcttgacccggtgtcaggcaacgtactccacatcctcacacacaaacgcct 1347
          ||| || ||  | ||||| |  || ||||||   ||| ||||||||| |||||||| |
Db    106 CCTCCCCGTCATCGACCCCGACTCGGGCAACACTCTCTACATCCTCACCCACAAACGCAT 165

Qy   1348 gctcaagttcctgcacatctttgtaagcctgggccca 1384
          |||||||||||  |  ||||| |      || ||||
Db    166 CCTCAAGTTCCTCAAACTCTTTATTGCAGAGGTCCCA 202
```

```
RESULT  11
AJ396118
LOCUS       AJ396118                758 bp    mRNA    linear   EST 25-JAN-2001
DEFINITION  AJ396118 dkfz426 Gallus gallus cDNA clone 25f16r1, mRNA sequence.
ACCESSION   AJ396118
VERSION     AJ396118.1  GI:7127728
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 758)
  AUTHORS   Abdrakhmanov,I., Lodygin,D., Geroth,P., Arakawa,H., Law,A., Plachy
            ,J., Korn,B. and Buerstedde,J.M.
  TITLE     A large database of chicken bursal ESTs as a resource for the
            analysis of vertebrate gene function
  JOURNAL   Genome Res. 10 (12), 2062-2069 (2000)
  MEDLINE   20568495
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinistr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES             Location/Qualifiers
     source          1. .758
                     /organism="Gallus gallus"
                     /strain="CB"
                     /db_xref="taxon:9031"
                     /clone="25f16r1"
                     /clone_lib="dkfz426"
                     /tissue_type="Bursa of Fabricius"
BASE COUNT      168 a    238 c    186 g    166 t
ORIGIN
```

```
  Query Match             4.6%;  Score 78.6;  DB 9;  Length 758;
  Best Local Similarity   68.8%;  Pred. No. 5e-07;
  Matches  108;  Conservative    0;  Mismatches   49;  Indels    0;  Gaps    0;

Qy   1228 ctcccatctctgcagcctgtttgaagctgtctacaccctcatcaagaaccggatccatcg 1287
          |||||       |||||| ||||| || |||| | |||| |||   ||   |||||| ||
Db    267 CTCCCCCAATGCCAGCCTTTTTGATGCCGTCTCCTCCCTGATCCGCAATAAGATCCACCG 326

Qy   1288 cctgcctgttcttgacccggtgtcaggcaacgtactccacatcctcacacacaaacgcct 1347
          ||| || ||  | ||||| |  || ||||||   ||| ||||||||| | |||||||
Db    327 CCTCCCCGTCATCGACCCCGACTCGGGCAACACTCTCTACATCCTCACCCCCAAACGCAT 386

Qy   1348 gctcaagttcctgcacatctttgtaagcctgggccca 1384
          |||||||||||  |  ||||| |      || ||||
Db    387 CCTCAAGTTCCTCAAACTCTTTATTGCAGAGGTCCCA 423
```

```
RESULT  12
AV603335
LOCUS       AV603335                576 bp    mRNA    linear   EST 27-NOV-2001
DEFINITION  AV603335 Bos taurus kidney fetus Bos taurus cDNA clone E1KI015F02
            5', mRNA sequence.
ACCESSION   AV603335
VERSION     AV603335.1  GI:9725661
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 576)
  AUTHORS   Takasuga,A., Hirotsune,S., Itoh,R., Jitohzono,A., Suzuki,H., Aso,H.
            and Sugimoto,Y.
  TITLE     Establishment of a high throughput EST sequencing system using
            poly(A) tail-removed cDNA libraries and determination of 36,000
            bovine ESTs
  JOURNAL   Nucleic Acids Res. 29 (22), E108 (2001)
  MEDLINE   21570554
COMMENT     Contact: Yoshikazu Sugimoto
            Animal Genetics Division
            Shirakawa Institute of Animal Genetics
            Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
            Tel: 81-248-25-5641
            Fax: 81-248-25-5725
            Email: kazusugi@cocoa.ocn.ne.jp
            Single pass sequencing.
            This clone was obtained from a polyA-deleted cDNA library.
FEATURES             Location/Qualifiers
     source          1. .576
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone="E1KI015F02"
                     /clone_lib="Bos taurus kidney fetus"
                     /tissue_type="kidney"
                     /dev_stage="fetus"
                     /lab_host="DH10B"
                     /note="Vector: pZL1; Site_1: SalI; Site_2: NotI; Poly A
                     was deleted from a NotI site"
BASE COUNT      144 a    150 c    122 g    158 t      2 others
ORIGIN
```

```
  Query Match             4.4%;  Score 76.4;  DB 9;  Length 576;
  Best Local Similarity   70.1%;  Pred. No. 1.4e-06;
```

```
  Matches  101;  Conservative    0;  Mismatches   43;  Indels    0;  Gaps    0;

Qy   1240 cagcctgtttgaagctgtctacaccctcatcaagaaccggatccatcgcctgcctgttct 1299
          |||| ||||||| |||||||   |  | ||    |||  ||||||  | |||| ||| |
Db    432 CAGCTTGTTTGATGCTGTCTCTTCATTAATTCGAAACAAGATCCACAGGCTGCCAGTTAT 491

Qy   1300 tgacccggtgtcaggcaacgtactccacatcctcacacacaaacgcctgctcaagttcct 1359
          ||||||||  |||||||||   ||  |||| ||||| ||||| ||| | ||||||||||
Db    492 TGACCCGGAATCAGGCAACACCCTGTACATNCTCACCCACAAGCGCATNCTCAAGTTCCT 551

Qy   1360 gcacatctttgtaagcctgggccc 1383
            |  | ||| | | |   |  |||
Db    552 CAAGTTATTTATCACCGAGTTCCC 575


RESULT  13
AV608257
LOCUS       AV608257                564 bp    mRNA    linear   EST 28-NOV-2001
DEFINITION  AV608257 Bos taurus kidney fetus Bos taurus cDNA clone E1KI045H03
            5', mRNA sequence.
ACCESSION   AV608257
VERSION     AV608257.1  GI:9738630
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 564)
  AUTHORS   Takasuga,A., Hirotsune,S., Itoh,R., Jitohzono,A., Suzuki,H., Aso,H.
            and Sugimoto,Y.
  TITLE     Establishment of a high throughput EST sequencing system using
            poly(A) tail-removed cDNA libraries and determination of 36,000
            bovine ESTs
  JOURNAL   Nucleic Acids Res. 29 (22), E108 (2001)
  MEDLINE   21570554
COMMENT     Contact: Yoshikazu Sugimoto
            Animal Genetics Division
            Shirakawa Institute of Animal Genetics
            Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8061, Japan
            Tel: 81-248-25-5641
            Fax: 81-248-25-5725
            Email: kazusugi@cocoa.ocn.ne.jp
            Single pass sequencing.
            This clone was obtained from a polyA-deleted cDNA library.
FEATURES             Location/Qualifiers
     source          1. .564
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone="E1KI045H03"
                     /clone_lib="Bos taurus kidney fetus"
                     /tissue_type="kidney"
                     /dev_stage="fetus"
                     /lab_host="DH10B"
                     /note="Vector: pZL1; Site_1: SalI; Site_2: NotI; Poly A
                     was deleted from a NotI site"
BASE COUNT      142 a    145 c    121 g    155 t      1 others
ORIGIN


  Query Match             4.2%;  Score 73;  DB 9;  Length 564;
  Best Local Similarity   72.3%;  Pred. No. 7.7e-06;
  Matches   94;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy   1240 cagcctgtttgaagctgtctacaccctcatcaagaaccggatccatcgcctgcctgttct 1299
          |||| ||||||| |||||||   |  | ||    |||  ||||||  | |||| ||| |
Db    434 CAGCTTGTTTGATGCTGTCTCTTCATTAATTCGAAACAAGATCCACAGGCTGCCAGTTAT 493

Qy   1300 tgacccggtgtcaggcaacgtactccacatcctcacacacaaacgcctgctcaagttcct 1359
          ||||||||  |||||||||   ||   |||||||||| ||||| ||| | ||||||||||
Db    494 TGACCCGGAATCAGGCAACACCCTGTNCATCCTCACCCACAAGCGCATCCTCAAGTTCCT 553

Qy   1360 gcacatcttt 1369
```

```
            |  | |||
Db    554 CAAGTTATTT 563


RESULT  14
BI833269
LOCUS       BI833269                775 bp    mRNA    linear   EST 04-OCT-2001
DEFINITION  603087149F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5226132 5',
            mRNA sequence.
ACCESSION   BI833269
VERSION     BI833269.1  GI:15944819
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 775)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11569  row: a  column: 13
            High quality sequence stop: 773.
FEATURES             Location/Qualifiers
     source          1. .775
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5226132"
                     /clone_lib="NIH_MGC_120"
                     /lab_host="DH10B"
                     /note="Organ: pooled pancreas and spleen; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of spleen and pancreas from 28 yo
                     male.  Library is oligo-dT primed and directionally cloned
                     (EcoRV  site is destroyed upon cloning).  Average insert
                     size 1.5 kb, insert size range 1-2.5 kb.  Library is
                     normalized and enriched for  full-length clones and was
                     constructed by C. Gruber (Invitrogen). Research Genetics
                     tracking code 025. Note: this is a NIH_MGC Library."
BASE COUNT      200 a    194 c    171 g    209 t      1 others
ORIGIN


  Query Match             4.1%;  Score 71;  DB 10;  Length 775;
  Best Local Similarity   68.1%;  Pred. No. 2.5e-05;
  Matches   98;  Conservative    0;  Mismatches   46;  Indels    0;  Gaps    0;

Qy   1240 cagcctgtttgaagctgtctacaccctcatcaagaaccggatccatcgcctgcctgttct 1299
          |||| ||||||| |||||||   |  | ||   ||||  ||||||  | |||| ||| |
Db    455 CAGCTTGTTTGATGCTGTCTCTTCATTAATTCGGAACAAGATCCACAGGCTGCCAGTTAT 514

Qy   1300 tgacccggtgtcaggcaacgtactccacatcctcacacacaaacgcctgctcaagttcct 1359
          ||||||  |  ||||||||     |  |||||||||| ||||| ||| | || ||||||||
Db    515 TGACCCAGAATCAGGCAATACTTTGTACATCCTCACCCACAAGCGCATTCTGAAGTTCCT 574

Qy   1360 gcacatctttgtaagcctgggccc 1383
            |  | ||| | |    |  |||
Db    575 CAAANTGTTTATCACTGAGTTCCC 598


RESULT  15
AA558845/c
LOCUS       AA558845                448 bp    mRNA    linear   EST 09-SEP-1997
DEFINITION  n169b09.s1 NCI_CGAP_Pr4.1 Homo sapiens cDNA clone IMAGE:1045913
            similar to SW:AAKG_RAT P80385 5'-AMP-ACTIVATED PROTEIN KINASE,
```

```
            GAMMA CHAIN ;, mRNA sequence.
ACCESSION   AA558845
VERSION     AA558845.1  GI:2329612
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 448)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuaqui, M.D.
            , Michael R. Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: David B. Krizman, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 561    Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 297.
FEATURES             Location/Qualifiers
     source          1. .448
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1045913"
                     /clone_lib="NCI_CGAP_Pr4.1"
                     /sex="male"
                     /tissue_type="prostatic intraepithelial neoplasia - high
                     grade"
                     /lab_host="DH10B"
                     /note="Organ: prostate; Vector: pAMP10; mRNA made from
                     prostatic intraepithelial neoplasia (high-grade), cDNA
                     made  by oligo-dT priming.  Non-directionally cloned.
                     Size-selected on agarose gel, average insert size 600 bp.
                     cDNA Library Preparation: David B. Krizman, Ph.D.
                     Reference:  Krizman et al. (1996) Cancer Research
                     56:5380-5383. cDNA Library Arrayed by: Greg Lennon, Ph.D.
                     DNA Sequencing by: Washington University Genome Sequencing
                     Center"
BASE COUNT      124 a     87 c    117 g    120 t
ORIGIN


  Query Match             4.1%;  Score 70.4;  DB 9;  Length 448;
  Best Local Similarity   68.1%;  Pred. No. 2.6e-05;
  Matches   98;  Conservative    0;  Mismatches   46;  Indels    0;  Gaps    0;

Qy    1240 cagcctgtttgaagctgtctacaccctcatcaagaaccggatccatcgcctgcctgttct 1299
           |||| ||||||| |||||||   |  | ||   ||||  ||||||  | ||||| ||| |
Db     238 CAGCTTGTTTGATGCTGTCTCTTCATTAATTCGGAACAAGATCCACAGGCTGCCAGTTAT 179

Qy    1300 tgacccggtgtcaggcaacgtactccacatcctcacacacaaacgcctgctcaagttcct 1359
           |||||| |  |||||||      |  |||||||||| ||||| ||| | || ||||||||
Db     178 TGACCCAGAATCAGGCAATACTTTGTACATCCTCACCCACAAGCGCATTCTGAAGTTCCT 119

Qy    1360 gcacatctttgtaagcctgggccc 1383
             |  | ||| | |    |  |||
Db     118 CAAATTGTTTATCACTGAGTTCCC 95


Search completed: October  3, 2002, 16:14:46
Job time: 16941 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 11:28:00 ; Search time 7316.32 Seconds
(without alignments)
4925.358 Million cell updates/sec

Title: US-09-826-581-3
Perfect score: 1722
Sequence: 1 cctggcccctcagatcaaga.....gatgagaggctcgggctgga 1722

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1722	100.0	1722	6	AX281580	AX281580 Sequence
c 2	1719.4	99.8	152129	2	AC027416	AC027416 Homo sapi
c 3	1697	98.5	196554	2	AC073128	AC073128 Homo sapi
c 4	1697	98.5	206854	9	AC009974	AC009974 Homo sapi
5	612.6	35.6	5888	4	AF214521	AF214521 Sus scrof
c 6	209.4	12.2	227724	2	AF336381	AF336381 Mus muscu
7	168.2	9.8	1647	6	AX281582	AX281582 Sequence
8	168.2	9.8	2109	6	AX099776	AX099776 Sequence
9	168.2	9.8	2115	6	AX099802	AX099802 Sequence
10	168.2	9.8	2115	9	AF214519	AF214519 Homo sapi
11	168.2	9.8	2290	9	HSA249977	AJ249977 Homo sapi
12	148	8.6	1867	6	AX099774	AX099774 Sequence
13	148	8.6	1873	4	AF214520	AF214520 Sus scrof
14	148	8.6	1873	6	AX099800	AX099800 Sequence
15	148	8.6	2022	6	AX099804	AX099804 Sequence
16	91.4	5.3	106	11	G67375	G67375 A898 Human
17	82.2	4.8	142903	2	AL627254	AL627254 Danio rer
18	79.4	4.6	14411	4	AF329081	AF329081 Bos tauru
c 19	77	4.5	73638	2	AC015613	AC015613 Homo sapi
c 20	74.2	4.3	81210	2	AC019242	AC019242 Homo sapi
21	74.2	4.3	171822	2	AC011603	AC011603 Homo sapi
22	74.2	4.3	180749	2	AC025256	AC025256 Homo sapi
23	70.4	4.1	1578	9	HSU42412	U42412 Human 5'-AM
24	70.4	4.1	1677	9	BC000358	BC000358 Homo sapi
25	69.4	4.0	101215	9	AC006966	AC006966 Homo sapi
26	68	3.9	3497	3	AF094763	AF094763 Drosophil
27	68	3.9	80069	2	AC019671	AC019671 Drosophil
28	68	3.9	173634	3	AC009344	AC009344 Drosophil
29	68	3.9	195868	3	AC008308	AC008308 Drosophil
30	68	3.9	230266	3	AE003733	AE003733 Drosophil
31	67.6	3.9	7218	6	I66494	I66494 Sequence 14
32	65.8	3.8	2082	3	AF094764	AF094764 Drosophil
33	65.8	3.8	4666	3	AY070541	AY070541 Drosophil
34	65.6	3.8	104	10	AF266698S2	AF266699 Mus muscu
35	65.6	3.8	1328	10	RNAMPKGAM	X95578 R.norvegicu
36	65.6	3.8	1550	10	RNU42413	U42413 Rattus norv
37	65.6	3.8	1623	10	AF036535	AF036535 Mus muscu
38	63.8	3.7	1167	9	AB025580	AB025580 Homo sapi
39	63.8	3.7	1435	6	AR139104	AR139104 Sequence
40	63.8	3.7	2062	9	HSA249976	AJ249976 Homo sapi
41	63.8	3.7	2194	9	AF087875	AF087875 Homo sapi
42	63.8	3.7	2203	9	BC020540	BC020540 Homo sapi
43	63.8	3.7	2223	9	AK001887	AK001887 Homo sapi
44	63.8	3.7	3132	10	BC015283	BC015283 Mus muscu
45	62.4	3.6	125020	9	AF429315	AF429315 Homo sapi

ALIGNMENTS

RESULT 1
AX281580
LOCUS AX281580 1722 bp DNA linear PAT 03-NOV-2001
DEFINITION Sequence 3 from Patent WO0177305.
ACCESSION AX281580
VERSION AX281580.1 GI:16608831
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Andersson,L., Luthman,H. and Marklund,S.
TITLE Variants of the human amp-activated protein kinase gamma 3 subunit
JOURNAL Patent: WO 0177305-A 3 18-OCT-2001;
Arexis AB (SE)
FEATURES Location/Qualifiers
source 1. .1722
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 321 a 504 c 534 g 363 t
ORIGIN

```
  Query Match             100.0%;  Score 1722;  DB 6;  Length 1722;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1722;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 cctggccccctcagatcaagaaggccttctttgctctggtggccaacggtgtgcgggcagc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CCTGGCCCCTCAGATCAAGAAGGCCTTCTTTGCTCTGGTGGCCAACGGTGTGCGGGCAGC 60

Qy      61 ccctctatgggacagcaagaagcagagctttgtgggtgaggagaggctggggaggtgaag 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 CCCTCTATGGGACAGCAAGAAGCAGAGCTTTGTGGGTGAGGAGAGGCTGGGGAGGTGAAG 120

Qy     121 ggagatggaggaggtgagggggagatcttgtacggttgttctggggctgatctctgatat 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GGAGATGGAGGAGGTGAGGGGGAGATCTTGTACGGTTGTTCTGGGGCTGATCTCTGATAT 180

Qy     181 accacaagcttggcttcaggccaagcccagccaggggccagggtggaggaaagtccatcc 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 ACCACAAGCTTGGCTTCAGGCCAAGCCCAGCCAGGGGCCAGGGTGGAGGAAAGTCCATCC 240

Qy     241 ggagtctgcatggccagctgggagaccctggggctcaatttccccatctgtggagccgct 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GGAGTCTGCATGGCCAGCTGGGAGACCCTGGGGCTCAATTTCCCCATCTGTGGAGCCGCT 300

Qy     301 atgaccagctgacacctttcacctccgctactgcatggccctgtgccataggtgctaggg 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ATGACCAGCTGACACCTTTCACCTCCGCTACTGCATGGCCCTGTGCCATAGGTGCTAGGG 360

Qy     361 agcaaatgggggggaggcaggagagaaagagccccacttctcaggcctgggggggctgcccc 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 AGCAAATGGGGGGAGGCAGGAGAGAAAGAGCCCCACTTCTCAGGCCTGGGGGGCTGCCCC 420

Qy     421 actgtcctgttcccacagtccccactgtgtctcagcacaaggacactggcagggtgggga 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 ACTGTCCTGTTCCCACAGTCCCCACTGTGTCTCAGCACAAGGACACTGGCAGGGTGGGGA 480

Qy     481 ggggatctgaccctcaacctgccttccacccaaaggccccgggctgacctcctccccgcc 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 GGGGATCTGACCCTCAACCTGCCTTCCACCCAAAGGCCCCGGGCTGACCTCCTCCCCGCC 540

Qy     541 cctcccctgcagggatgctgaccatcactgacttcatcctggtgctgcatcgctactaca 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 CCTCCCCTGCAGGGATGCTGACCATCACTGACTTCATCCTGGTGCTGCATCGCTACTACA 600

Qy     601 ggtccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcgg 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GGTCCCCCCTGGTGAGGAGTGGGCTGGGAATCTTATGGGCACCCAGAGGGGCGGGGGCGG 660

Qy     661 aggggagtcctcctggagcctggtgccctagaagcccacgtctttctgacttctggagtc 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 AGGGGAGTCCTCCTGGAGCCTGGTGCCCTAGAAGCCCACGTCTTTCTGACTTCTGGAGTC 720

Qy     721 ctgtcgatgtctctaggtccagatctatgagattgaacaacataagattgagacctggag 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 CTGTCGATGTCTCTAGGTCCAGATCTATGAGATTGAACAACATAAGATTGAGACCTGGAG 780

Qy     781 gggtgagtggggagaggaacccggaaaggggctgttggtgatggtgggccagggcttaag 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 GGGTGAGTGGGGAGAGGAACCCGGAAAGGGGCTGTTGGTGATGGTGGGCCAGGGCTTAAG 840

Qy     841 gtggaggatgggcagtggggatgtcctggagtgaacaggggagggacaataggagcctcg 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 GTGGAGGATGGGCAGTGGGGATGTCCTGGAGTGAACAGGGGAGGGACAATAGGAGCCTCG 900

Qy     901 ggtgcctgacggaagggaagctgcctgggactgcaaggtgaggcaggtgaccggctcccc 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GGTGCCTGACGGAAGGGAAGCTGCCTGGGACTGCAAGGTGAGGCAGGTGACCGGCTCCCC 960

Qy     961 tggcctgactctggctctttctgcagagatctacctgcaaggctgcttcaagcctctggt 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 TGGCCTGACTCTGGCTCTTTCTGCAGAGATCTACCTGCAAGGCTGCTTCAAGCCTCTGGT 1020

Qy    1021 ctccatctctcctaatgataggtgggtgtctctgctcattcacctgagcctcctcctccc 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 CTCCATCTCTCCTAATGATAGGTGGGTGTCTCTGCTCATTCACCTGAGCCTCCTCCTCCC 1080

Qy    1081 acagtccccttccccagtcccactcagctctgaactcacctcttcatcctaggcggcaca 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 ACAGTCCCCTTCCCCAGTCCCACTCAGCTCTGAACTCACCTCTTCATCCTAGGCGGCACA 1140

Qy    1141 cagacaagggagccttggtgccctgccctccttttttaggggcctgggatggaggttgtct 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 CAGACAAGGGAGCCTTGGTGCCCTGCCCTCCTTTTTAGGGGCCTGGGATGGAGGTTGTCT 1200

Qy    1201 ctccctaggctgccccgaggctcactgctcccatctctgcagcctgtttgaagctgtcta 1260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 CTCCCTAGGCTGCCCCGAGGCTCACTGCTCCCATCTCTGCAGCCTGTTTGAAGCTGTCTA 1260

Qy    1261 caccctcatcaagaaccggatccatcgcctgcctgttcttgacccggtgtcaggcaacgt 1320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 CACCCTCATCAAGAACCGGATCCATCGCCTGCCTGTTCTTGACCCGGTGTCAGGCAACGT 1320

Qy    1321 actccacatcctcacacacaaacgcctgctcaagttcctgcacatctttgtaagcctggg 1380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 ACTCCACATCCTCACACACAAACGCCTGCTCAAGTTCCTGCACATCTTTGTAAGCCTGGG 1380

Qy    1381 cccaggtgggaggaaggggggagacctgggcaggtgatcagagggcctgaggagtcttcag 1440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 CCCAGGTGGGAGGAAGGGGGAGACCTGGGCAGGTGATCAGAGGGCCTGAGGAGTCTTCAG 1440

Qy    1441 ccctagcagtcgtggggaagagctgggagccctcttgaagctgctggatccctgatctcc 1500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 CCCTAGCAGTCGTGGGGAAGAGCTGGGAGCCCTCTTGAAGCTGCTGGATCCCTGATCTCC 1500

Qy    1501 acctggtccccatcctaaccagggttccctgctgccccggccctccttcctctaccgcac 1560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 ACCTGGTCCCCATCCTAACCAGGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCAC 1560

Qy    1561 tatccaagatttgggcatcggcacattccgagacttggctgtggtgctggagacagcacc 1620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 TATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCTGTGGTGCTGGAGACAGCACC 1620

Qy    1621 catcctgactgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaa 1680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 CATCCTGACTGCACTGGACATCTTTGTGGACCGGCGTGTGTCTGCACTGCCTGTGGTCAA 1680

Qy    1681 cgaatgtggtacccacccccaggatgagaggctcgggctgga 1722
           |||||||||||||||||||||||||||||||||||||||||||
Db    1681 CGAATGTGGTACCCACCCCCAGGATGAGAGGCTCGGGCTGGA 1722


RESULT   2
AC027416/c
LOCUS       AC027416              152129 bp    DNA     linear   HTG 07-JUN-2000
DEFINITION  Homo sapiens clone RP11-504G11, WORKING DRAFT SEQUENCE, 32
            unordered pieces.
ACCESSION   AC027416
VERSION     AC027416.2  GI:8317289
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 152129)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-504G11
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 152129)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
```

Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE Direct Submission

JOURNAL Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT On Jun 7, 2000 this sequence version replaced gi:7342115.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-------------- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-------------- Project Information
Center project name: L7458
Center clone name: 504_G_11
-------------- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 135376 bases at least Q40
Consensus quality: 143264 bases at least Q30
Consensus quality: 146503 bases at least Q20
Insert size: 161000; agarose-fp
Insert size: 149029; sum-of-contigs
Quality coverage: 3.1 in Q20 bases; agarose-fp
Quality coverage: 3.3 in Q20 bases; sum-of-contigs
--------------.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 32 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1005: contig of 1005 bp in length
* 1006 1105: gap of 100 bp
* 1106 2402: contig of 1297 bp in length
* 2403 2502: gap of 100 bp
* 2503 3823: contig of 1321 bp in length
* 3824 3923: gap of 100 bp
* 3924 5020: contig of 1097 bp in length
* 5021 5120: gap of 100 bp
* 5121 6161: contig of 1041 bp in length
* 6162 6261: gap of 100 bp
* 6262 7547: contig of 1286 bp in length
* 7548 7647: gap of 100 bp
* 7648 9983: contig of 2336 bp in length
* 9984 10083: gap of 100 bp
* 10084 12556: contig of 2473 bp in length
* 12557 12656: gap of 100 bp
* 12657 15043: contig of 2387 bp in length
* 15044 15143: gap of 100 bp
* 15144 17123: contig of 1980 bp in length
* 17124 17223: gap of 100 bp
* 17224 19466: contig of 2243 bp in length
* 19467 19566: gap of 100 bp
* 19567 21928: contig of 2362 bp in length
* 21929 22028: gap of 100 bp
* 22029 24319: contig of 2291 bp in length
* 24320 24419: gap of 100 bp
* 24420 27059: contig of 2640 bp in length
* 27060 27159: gap of 100 bp
* 27160 30170: contig of 3011 bp in length
* 30171 30270: gap of 100 bp
* 30271 33968: contig of 3698 bp in length
* 33969 34068: gap of 100 bp
* 34069 38179: contig of 4111 bp in length
* 38180 38279: gap of 100 bp
* 38280 42366: contig of 4087 bp in length
* 42367 42466: gap of 100 bp
* 42467 46365: contig of 3899 bp in length
* 46366 46465: gap of 100 bp
* 46466 51285: contig of 4820 bp in length
* 51286 51385: gap of 100 bp
* 51386 55871: contig of 4486 bp in length
* 55872 55971: gap of 100 bp
* 55972 60595: contig of 4624 bp in length
* 60596 60695: gap of 100 bp
* 60696 66595: contig of 5900 bp in length
* 66596 66695: gap of 100 bp
* 66696 73218: contig of 6523 bp in length
* 73219 73318: gap of 100 bp
* 73319 77115: contig of 3797 bp in length
* 77116 77215: gap of 100 bp
* 77216 85022: contig of 7807 bp in length
* 85023 85122: gap of 100 bp
* 85123 93314: contig of 8192 bp in length
* 93315 93414: gap of 100 bp
* 93415 101193: contig of 7779 bp in length
* 101194 101293: gap of 100 bp
* 101294 113090: contig of 11797 bp in length
* 113091 113190: gap of 100 bp
* 113191 123496: contig of 10306 bp in length
* 123497 123596: gap of 100 bp
* 123597 137837: contig of 14241 bp in length
* 137838 137937: gap of 100 bp
* 137938 152129: contig of 14192 bp in length.

FEATURES Location/Qualifiers

source 1. .152129
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-504G11"
/clone_lib="RPCI-11 Human Male BAC"

misc_feature 1. .1005
/note="assembly_fragment"

misc_feature 1106. .2402
/note="assembly_fragment"

misc_feature 2503. .3823
/note="assembly_fragment"

misc_feature 3924. .5020
/note="assembly_fragment"

misc_feature 5121. .6161
/note="assembly_fragment"

misc_feature 6262. .7547
/note="assembly_fragment"

misc_feature 7648. .9983
/note="assembly_fragment"

misc_feature 10084. .12556
/note="assembly_fragment"

misc_feature 12657. .15043
/note="assembly_fragment"

misc_feature 15144. .17123
/note="assembly_fragment"

misc_feature 17224. .19466
/note="assembly_fragment
clone_end:SP6
vector_side:left"

misc_feature 19567. .21928
/note="assembly_fragment"

misc_feature 22029. .24319
/note="assembly_fragment"

misc_feature 24420. .27059
/note="assembly_fragment"
misc_feature 27160. .30170
/note="assembly_fragment"
misc_feature 30271. .33968
/note="assembly_fragment"
misc_feature 34069. .38179
/note="assembly_fragment"
misc_feature 38280. .42366
/note="assembly_fragment"
misc_feature 42467. .46365
/note="assembly_fragment"
misc_feature 46466. .51285
/note="assembly_fragment"
misc_feature 51386. .55871
/note="assembly_fragment"
misc_feature 55972. .60595
/note="assembly_fragment"
misc_feature 60696. .66595
/note="assembly_fragment"
misc_feature 66696. .73218
/note="assembly_fragment"
misc_feature 73319. .77115
/note="assembly_fragment"
misc_feature 77216. .85022
/note="assembly_fragment"
misc_feature 85123. .93314
/note="assembly_fragment"
misc_feature 93415. .101193
/note="assembly_fragment"
misc_feature 101294. .113090
/note="assembly_fragment"

Query Match 99.8%; Score 1719.4; DB 2; Length 152129;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1720; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy       1 cctggccccctcagatcaagaaggccttctttgctctggtggccaacggtgtgcgggcagc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   36618 CCTGGCCCCTCAGATCAAGAAGGCCTTCTTTGCTCTGGTGGCCAACGGTGTGCGGGCAGC 36559

Qy      61 ccctctatgggacagcaagaagcagagctttgtgggtgaggagaggctggggaggtgaag 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   36558 CCCTCTATGGGACAGCAAGAAGCAGAGCTTTGTGGGTGAGGAGAGGCTGGGGAGGTGAAG 36499

Qy     121 ggagatggaggaggtgaggggagatcttgtacggttgttctggggctgatctctgatat 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   36498 GGAGATGGAGGAGGTGAGGGGGAGATCTTGTACGGTTGTTCTGGGGCTGATCTCTGATAT 36439

Qy     181 accacaagcttggcttcaggccaagcccagccaggggccagggtggaggaaagtccatcc 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   36438 ACCACAAGCTTGGCTTCAGGCCAAGCCCAGCCAGGGGCCAGGGTGGAGGAAAGTCCATCC 36379

Qy     241 ggagtctgcatggccagctgggagaccctggggctcaatttccccatctgtggagccgct 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   36378 GGAGTCTGCATGGCCAGCTGGGAGACCCTGGGGCTCAATTTCCCCATCTGTGGAGCCGCT 36319

Qy     301 atgaccagctgacacctttcacctccgctactgcatggccctgtgccataggtgctaggg 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   36318 ATGACCAGCTGACACCTTTCACCTCCGCTACTGCATGGCCCTGTGCCATAGGTGCTAGGG 36259

Qy     361 agcaaatgggggggaggcaggagagaaaagagccccacttctcaggcctgggggggctgcccc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   36258 AGCAAATGGGGGGAGGCAGGAGAGAAAGAGCCCCACTTCTCAGGCCTGGGGGGCTGCCCC 36199

Qy     421 actgtcctgttcccacagtccccactgtgtctcagcacaaggacactggcagggtgggga 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   36198 ACTGTCCTGTTCCCACAGTCCCCACTGTGTCTCAGCACAAGGACACTGGCAGGGTGGGGA 36139

Qy     481 ggggatctgaccctcaacctgccttccacccaaaggccccgggctgacctcctccccgcc 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   36138 GGGGATCTGACCCTCAACCTGCCTTCCACCCAAAGGCCCCGGGCTGACCTCCTCCCCGCC 36079

Qy     541 cctcccctgcagggatgctgaccatcactgacttcatcctggtgctgcatcgctactaca 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   36078 CCTCCCCTGCAGGGATGCTGACCATCACTGACTTCATCCTGGTGCTGCATCGCTACTACA 36019

Qy     601 ggtccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcgg 660
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db   36018 GGTCCCCCCTGGTGAGGAGTGGGCTGGGAATCTTATGGGCAGCCAGAGGGGCGGGGGCGG 35959

Qy     661 aggggagtcctcctggagcctggtgccctagaagcccacgtctttctgacttctggagtc 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35958 AGGGGAGTCCTCCTGGAGCCTGGTGCCCTAGAAGCCCACGTCTTTCTGACTTCTGGAGTC 35899

Qy     721 ctgtcgatgtctctaggtccagatctatgagattgaacaacataagattgagacctggag 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35898 CTGTCGATGTCTCTAGGTCCAGATCTATGAGATTGAACAACATAAGATTGAGACCTGGAG 35839

Qy     781 gggtgagtggggagaggaacccggaaaaggggctgttggtgatggtgggccagggcttaag 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35838 GGGTGAGTGGGGAGAGGAACCCGGAAAGGGGCTGTTGGTGATGGTGGGCCAGGGCTTAAG 35779

Qy     841 gtggaggatgggcagtggggatgtcctggagtgaacaggggagggacaataggagcctcg 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35778 GTGGAGGATGGGCAGTGGGGATGTCCTGGAGTGAACAGGGGAGGGACAATAGGAGCCTCG 35719

Qy     901 ggtgcctgacggaagggaagctgcctgggactgcaaggtgaggcaggtgaccggctcccc 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35718 GGTGCCTGACGGAAGGGAAGCTGCCTGGGACTGCAAGGTGAGGCAGGTGACCGGCTCCCC 35659

Qy     961 tggcctgactctggctctttctgcagagatctacctgcaaggctgcttcaagcctctggt 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35658 TGGCCTGACTCTGGCTCTTTCTGCAGAGATCTACCTGCAAGGCTGCTTCAAGCCTCTGGT 35599

Qy    1021 ctccatctctcctaatgataggtgggtgtctctgctcattcacctgagcctcctcctccc 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35598 CTCCATCTCTCCTAATGATAGGTGGGTGTCTCTGCTCATTCACCTGAGCCTCCTCCTCCC 35539

Qy    1081 acagtccccttccccagtcccactcagctctgaactcacctcttcatcctaggcggcaca 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35538 ACAGTCCCCTTCCCCAGTCCCACTCAGCTCTGAACTCACCTCTTCATCCTAGGCGGCACA 35479

Qy    1141 cagacaagggagccttggtgccctgccctccttttttaggggcctgggatggaggttgtct 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35478 CAGACAAGGGAGCCTTGGTGCCCTGCCCTCCTTTTTAGGGGCCTGGGATGGAGGTTGTCT 35419

Qy    1201 ctccctaggctgccccgaggctcactgctcccatctctgcagcctgtttgaagctgtcta 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35418 CTCCCTAGGCTGCCCCGAGGCTCACTGCTCCCATCTCTGCAGCCTGTTTGAAGCTGTCTA 35359

Qy    1261 caccctcatcaagaaccggatccatcgcctgcctgttcttgacccggtgtcaggcaacgt 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35358 CACCCTCATCAAGAACCGGATCCATCGCCTGCCTGTTCTTGACCCGGTGTCAGGCAACGT 35299

Qy    1321 actccacatcctcacacacaaacgcctgctcaagttcctgcacatctttgtaagcctggg 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35298 ACTCCACATCCTCACACACAAACGCCTGCTCAAGTTCCTGCACATCTTTGTAAGCCTGGG 35239

Qy    1381 cccaggtgggaggaaggggagacctgggcaggtgatcagagggcctgaggagtcttcag 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35238 CCCAGGTGGGAGGAAGGGGAGACCTGGGCAGGTGATCAGAGGGCCTGAGGAGTCTTCAG 35179

Qy    1441 ccctagcagtcgtggggaagagctgggagccctcttgaagctgctggatccctgatctcc 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35178 CCCTAGCAGTCGTGGGGAAGAGCTGGGAGCCCTCTTGAAGCTGCTGGATCCCTGATCTCC 35119

Qy    1501 acctggtccccatcctaaccagggttccctgctgccccggccctccttcctctaccgcac 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35118 ACCTGGTCCCCATCCTAACCAGGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCAC 35059

Qy    1561 tatccaagatttgggcatcggcacattccgagacttggctgtggtgctggagacagcacc 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   35058 TATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCTGTGGTGCTGGAGACAGCACC 34999

Qy    1621 catcctgactgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaa 1680
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  34998 CATCCTGACTGCACTGGACATCTTTGTGGACCGGCGTGTGTCTGCACTGCCTGTGGTCAA 34939

Qy   1681 cqaatgtggtacccacccccaqgatgagaggctcgggctgg 1721
          |||||||||||||||||||||||||||||||||||||||||
Db  34938 CGAATGTGGTACCCACCCCCAGGATGAGAGGCTCGGGCTGG 34898


RESULT  3
AC073128/c
LOCUS       AC073128              196554 bp    DNA     linear   HTG 21-FEB-2001
DEFINITION  Homo sapiens chromosome 2 clone RP11-647O5, WORKING DRAFT SEQUENCE,
            17 unordered pieces.
ACCESSION   AC073128
VERSION     AC073128.3  GI:13027579
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 196554)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 196554)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JUN-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Feb 21, 2001 this sequence version replaced gi:8469048.

            ----------------- Genome Center -----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: H_NH0647O05
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 98%
            Sequencing vector: plasmid; 0%
            Chemistry: Dye-primer ET; 98% of reads
            Chemistry: Dye-terminator Big Dye; 0% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 187795 bases at least Q40
            Consensus quality: 190513 bases at least Q30
            Consensus quality: 192099 bases at least Q20
            Insert size: 200000; agarose-fp
            Insert size: 194954; sum-of-contigs
            Quality coverage: 5.58 in Q20 bases; agarose-fp
            Quality coverage: 5.67 in Q20 bases; sum-of-contigs
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 17 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1157: contig of 1157 bp in length
            *     1158     1257: gap of unknown length
            *     1258     3600: contig of 2343 bp in length
            *     3601     3700: gap of unknown length
            *     3701     5103: contig of 1403 bp in length
            *     5104     5203: gap of unknown length
            *     5204     8524: contig of 3321 bp in length
            *     8525     8624: gap of unknown length
            *     8625    11856: contig of 3232 bp in length
            *    11857    11956: gap of unknown length
            *    11957    15783: contig of 3827 bp in length
            *    15784    15883: gap of unknown length
            *    15884    21906: contig of 6023 bp in length
            *    21907    22006: gap of unknown length
            *    22007    28887: contig of 6881 bp in length
            *    28888    28987: gap of unknown length
            *    28988    35255: contig of 6268 bp in length
            *    35256    35355: gap of unknown length
            *    35356    44642: contig of 9287 bp in length
            *    44643    44742: gap of unknown length
            *    44743    58275: contig of 13533 bp in length
            *    58276    58375: gap of unknown length
            *    58376    73816: contig of 15441 bp in length
            *    73817    73916: gap of unknown length
            *    73917    92140: contig of 18224 bp in length
            *    92141    92240: gap of unknown length
            *    92241   113337: contig of 21097 bp in length
            *   113338   113437: gap of unknown length
            *   113438   130325: contig of 16888 bp in length
            *   130326   130425: gap of unknown length
            *   130426   149287: contig of 18862 bp in length
            *   149288   149387: gap of unknown length
            *   149388   196554: contig of 47167 bp in length.
FEATURES             Location/Qualifiers
     source          1. .196554
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /clone="RP11-647O5"
     misc_feature    1. .1157
                     /note="assembly_name:Contig17"
     misc_feature    1258. .3600
                     /note="assembly_name:Contig18"
     misc_feature    3701. .5103
                     /note="assembly_name:Contig19
                     clone_end:T7
                     vector_side:right"
     misc_feature    5204. .8524
                     /note="assembly_name:Contig20"
     misc_feature    8625. .11856
                     /note="assembly_name:Contig21"
     misc_feature    11957. .15783
                     /note="assembly_name:Contig22"
     misc_feature    15884. .21906
                     /note="assembly_name:Contig23"
     misc_feature    22007. .28887
                     /note="assembly_name:Contig24"
     misc_feature    28988. .35255
                     /note="assembly_name:Contig25"
     misc_feature    35356. .44642
                     /note="assembly_name:Contig26"
     misc_feature    44743. .58275
                     /note="assembly_name:Contig27"
     misc_feature    58376. .73816
                     /note="assembly_name:Contig28"
     misc_feature    73917. .92140
                     /note="assembly_name:Contig29"
     misc_feature    92241. .113337
                     /note="assembly_name:Contig30"
     misc_feature    113438. .130325
                     /note="assembly_name:Contig31"
     misc_feature    130426. .149287
                     /note="assembly_name:Contig32"
     misc_feature    149388. .196554
                     /note="assembly_name:Contig33
                     clone_end:SP6
                     vector_side:right"
BASE COUNT    52296 a  46993 c  45889 g  49770 t   1606 others
ORIGIN


  Query Match             98.5%;  Score 1697;  DB 2;  Length 196554;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1719;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    2;

Qy      1 cctggcccctcagatcaagaaggccttctttgctctggtggccaacggtgtgcgggcagc 60
```

```
Db  61117 CCTGGCCCCTCAGATCAAGAAGGCCTTCTTTGCTCTGGTGGCCAACGGTGTGCGGGCAGC 61058

Qy     61 ccctctatgggacagcaagaagcagagctttgtgggtgaggagaggctggggaggtgaag 120
Db  61057 CCCTCTATGGGACAGCAAGAAGCAGAGCTTTGTGGGTGAGGAGAGGCTGGGGAGGTGAAG 60998

Qy    121 ggagatggaggaggtgaggggggagatcttgtacggttgttctggggctgatctctgatat 180
Db  60997 GGAGATGGAGGAGGTGAGGGGGAGATCTTGTACGGTTGTTCTGGGGCTGATCTCTGATAT 60938

Qy    181 accacaagcttggcttcaggccaagcccagccagggccagggtggaggaaagtccatcc 240
Db  60937 ACCACAAGCTTGGCTTCAGGCCAAGCCCAGCCAGGGGCCAGGGTGGAGGAAAGTCCATCC 60878

Qy    241 ggagtctgcatggccagctgggagaccctggggctcaatttccccatctgtggagccgct 300
Db  60877 GGAGTCTGCATGGCCAGCTGGGAGACCCTGGGGCTCAATTTCCCCATCTGTGGAGCCGCT 60818

Qy    301 atgaccagctgacacctttcacctccgctactgcatggccctgtgccataggtgctaggg 360
Db  60817 ATGACCAGCTGACACCTTTCACCTCCGCTACTGCATGGCCCTGTGCCATAGGTGCTAGGG 60758

Qy    361 agcaaatgggggaggcaggagagaaagagccccacttctcaggcctgggggctgcccc 420
Db  60757 AGCAAATGGGGGAGGCAGGAGAGAAAGAGCCCCACTTCTCAGGCCTGGGGGGCTGCCCC 60698

Qy    421 actgtcctgttcccacagtccccactgtgtctcagcacaaggacactggcagggtgggga 480
Db  60697 ACTGTCCTGTTCCCACAGTCCCCACTGTGTCTCAGCACAAGGACACTGGCAGGGTGGGGA 60638

Qy    481 ggggatctgaccctcaacctgccttccacccaaaggccccgggctgacctcctccccgcc 540
Db  60637 GGGGATCTGACCCTCAACCTGCCTTCCACCCAAAGGCCCCGGGCTGACCTCCTCCCCGCC 60578

Qy    541 cctcccctgcagggatgctgaccatcactgacttcatcctggtgctgcatcgctactaca 600
Db  60577 CCT-CCCTGCAGGGATGCTGACCATCACTGACTTCATCCTGGTGCTGCATCGCTACTACA 60519

Qy    601 ggtcccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcgg 660
Db  60518 GGTCCCCCCTGGTGAGGAGTGGGCTGGGAATCTTATGGGCACCCAGAGGGGCGGGGGCGG 60459

Qy    661 aggggagtcctcctggagcctggtgccctagaagcccacgtctttctgacttctggagtc 720
Db  60458 AGGGGAGTCCTCCTGGAGCCTGGTGCCCTAGAAGCCCACGTCTTTCTGACTTCTGGAGTC 60399

Qy    721 ctgtcgatgtctctaggtccagatctatgagattgaacaacataagattgagacctggag 780
Db  60398 CTGTCGATGTCTCTAGGTCCAGATCTATGAGATTGAACAACATAAGATTGAGACCTGGAG 60339

Qy    781 gggtgagtggggagaggaacccggaaaggggctgttggtgatggtgggccagggcttaag 840
Db  60338 GGGTGAGTGGGGAGAGGAACCCGGAAA-GGGCTGTTGGTGATGGTGGGCCAGGGCTTAAG 60280

Qy    841 gtggaggatgggcagtggggatgtcctggagtgaacaggggagggacaataggagcctcg 900
Db  60279 GTGGAGGATGGGCAGTGGGGATGTCCTGGAGTGAACAGGGGAGGGACAATAGGAGCCTCG 60220

Qy    901 ggtgcctgacggaagggaagctgcctgggactgcaaggtgaggcaggtgaccggctcccc 960
Db  60219 GGTGCCTGACGGAAGGGAAGCTGCCTGGGACTGCAAGGTGAGGCAGGTGACCGGCTCCCC 60160

Qy    961 tggcctgactctggctctttctgcagagatctacctgcaaggctgcttcaagcctctggt 1020
Db  60159 TGGCCTGACTCTGGCTCTTTCTGCAGAGATCTACCTGCAAGGCTGCTTCAAGCCTCTGGT 60100

Qy   1021 ctccatctctcctaatgataggtgggtgtctctgctcattcacctgagcctcctcctccc 1080
Db  60099 CTCCATCTCTCCTAATGATAGGTGGGTGTCTCTGCTCATTCACCTGAGCCTCCTCCTCCC 60040

Qy   1081 acagtcccctccccagtcccactcagctctgaactcacctcttcatcctaggcggcaca 1140
Db  60039 ACAGTCCCCTTCCCCAGTCCCACTCAGCTCTGAACTCACCTCTTCATCCTAGGCGGCACA 59980

Qy   1141 cagacaagggagccttggtgccctgccctccttttttaggggcctgggatggaggttgtct 1200
Db  59979 CAGACAAGGGAGCCTTGGTGCCCTGCCCTCCTTTTTAGGGGCCTGGGATGGAGGTTGTCT 59920

Qy   1201 ctccctaggctgccccgaggctcactgctcccatctctgcagcctgtttgaagctgtcta 1260
Db  59919 CTCCCTAGGCTGCCCCGAGGCTCACTGCTCCCATCTCTGCAGCCTGTTTGAAGCTGTCTA 59860

Qy   1261 caccctcatcaagaaccggatccatcgcctgcctgttcttgacccggtgtcaggcaacgt 1320
Db  59859 CACCCTCATCAAGAACCGGATCCATCGCCTGCCTGTTCTTGACCCGGTGTCAGGCAACGT 59800

Qy   1321 actccacatcctcacacacaaacgcctgctcaagttcctgcacatctttgtaagcctggg 1380
Db  59799 ACTCCACATCCTCACACACAAACGCCTGCTCAAGTTCCTGCACATCTTTGTAAGCCTGGG 59740

Qy   1381 cccaggtgggaggaaggggagacctgggcaggtgatcagagggcctgaggagtcttcag 1440
Db  59739 CCCAGGTGGGAGGAAGGGGGAGACCTGGGCAGGTGATCAGAGGGCCTGAGGAGTCTTCAG 59680

Qy   1441 ccctagcagtcgtggggaagagctgggagccctcttgaagctgctggatccctgatctcc 1500
Db  59679 CCCTAGCAGTCGTGGGGAAGAGCTGGGAGCCCTCTTGAAGCTGCTGGATCCCTGATCTCC 59620

Qy   1501 acctggtccccatcctaaccagggttccctgctgccccggcccctccttcctctaccgcac 1560
Db  59619 ACCTGGTCCCCATCCTAACCAGGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCAC 59560

Qy   1561 tatccaagatttgggcatcggcacattccgagacttggctgtggtgctggagacagcacc 1620
Db  59559 TATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCTGTGGTGCTGGAGACAGCACC 59500

Qy   1621 catcctgactgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaa 1680
Db  59499 CATCCTGACTGCACTGGACATCTTTGTGGACCGGCGTGTGTCTGCACTGCCTGTGGTCAA 59440

Qy   1681 cgaatgtggtacccacccccaggatgagaggctcgggctgg 1721
Db  59439 CGAATGTGGTACCCACCCCCAGGATGAGAGGCTCGGGCTGG 59399
```

```
RESULT   4
AC009974/c
LOCUS       AC009974              206854 bp    DNA     linear   PRI 09-JAN-2002
DEFINITION  Homo sapiens BAC clone RP11-459I19 from 2, complete sequence.
ACCESSION   AC009974
VERSION     AC009974.9  GI:16799058
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 206854)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
REFERENCE   2  (bases 1 to 206854)
  AUTHORS   Harris,A. and Cotton,M.
  TITLE     The sequence of Homo sapiens BAC clone RP11-459I19
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 206854)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 206854)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-NOV-2001) Genome Sequencing Center, Washington
```

```
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 206854)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JAN-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   6  (bases 1 to 206854)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JAN-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Nov 8, 2001 this sequence version replaced gi:13431203.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0459I19
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. John D.
            McPherson, Department of Genetics, Washington University, St. Louis
            MO.  For additional information about the map position of this
            sequence, see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries.  Genomics 51:1-8.  The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at the Roswell Park Cancer Institute
            (http://bacpac.med.buffalo.edu)
            VECTOR:  pBACe3.6
            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is RP11-1077K22; the clone
            sequenced to the right is RP11-647O5.  Actual start of this clone
            is at base position 1 of RP11-459I19; actual end is at base
            position 206854 of RP11-459I19.

            Data from AC079810 and AC073128 was used to finish this clone,
            AC009974.  Polymorphisms have been identified between AC073128 and
            AC009974.  A single plasmid region exists beteen 38812-38903.  Am
            unresolved tandem in the HERV SVA exists between 184390-185163.
            PCR suggests that approximately 1700 bps are missing.
FEATURES            Location/Qualifiers
     source         1. .206854
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /chromosome="2"
                    /map="2"
                    /clone="RP11-459I19"
                    /clone_lib="RPCI-11"
     misc_feature   1. .37
                    /note="match to EST BE314060 (NID:g9134719)"
     misc_feature   1. .37
                    /note="match to EST BI059713 (NID:g14467240)"
     misc_feature   1. .37
                    /note="match to EST BF183086 (NID:g11061273)"
     misc_feature   1. .37
                    /note="match to EST AL567345 (NID:g12920610)"
     misc_feature   1. .37
                    /note="match to EST AW880850 (NID:g8042860)"
     misc_feature   1. .37
                    /note="match to EST BF304755 (NID:g11251653)"
     misc_feature   1. .37
                    /note="similar to Homo sapiens EST BI114348
                    (NID:g14565249)"
     misc_feature   1. .37
                    /note="match to EST BG477625 (NID:g13409904)"
     misc_feature   1. .37
                    /note="match to EST BE047599 (NID:g8364652) tz39c01.y1"
     misc_feature   3. .37
                    /note="match to EST BE908408 (NID:g10402954)"
     misc_feature   4. .37
                    /note="match to EST AI670836 (NID:g4850567) wa04g10.x1"
     misc_feature   164. .662
                    /note="match to EST C05773 (NID:g1502549)"
     misc_feature   220. .221
                    /note="match to EST BG470047 (NID:g13402322)"
     misc_feature   281. .344
                    /note="match to EST BE047599 (NID:g8364652) tz39c01.y1"
     misc_feature   281. .344
                    /note="match to EST BE908408 (NID:g10402954)"
     misc_feature   281. .344
                    /note="match to EST BF304755 (NID:g11251653)"
     misc_feature   281. .344
                    /note="match to EST BF183086 (NID:g11061273)"
     misc_feature   281. .344
                    /note="similar to Homo sapiens EST BI114348
                    (NID:g14565249)"
     misc_feature   281. .344
                    /note="match to EST BG477625 (NID:g13409904)"
     misc_feature   281. .344
                    /note="match to EST BI059713 (NID:g14467240)"
     misc_feature   281. .344
                    /note="match to EST AL567345 (NID:g12920610)"
     misc_feature   281. .344
                    /note="match to EST BE314060 (NID:g9134719)"
     misc_feature   281. .344
                    /note="match to EST AI670836 (NID:g4850567) wa04g10.x1"
     misc_feature   281. .344
                    /note="match to EST AW880850 (NID:g8042860)"
     misc_feature   294. .344
                    /note="match to EST BG470047 (NID:g13402322)"
     misc_feature   complement(557. .558)
                    /note="match to EST AA043371 (NID:g1521226) zk53e10.r1"
     misc_feature   588. .929
                    /note="match to EST T64073 (NID:g667938) yc05d12.r1"
     misc_feature   594. .763
                    /note="match to EST BF304755 (NID:g11251653)"
     misc_feature   594. .763
                    /note="match to EST BG470047 (NID:g13402322)"
     misc_feature   594. .763
                    /note="match to EST AW880850 (NID:g8042860)"
     misc_feature   594. .731
                    /note="match to EST BE314060 (NID:g9134719)"
     misc_feature   594. .764
                    /note="match to EST BI059713 (NID:g14467240)"
     misc_feature   594. .763
                    /note="match to EST BG477625 (NID:g13409904)"
     misc_feature   594. .742
                    /note="match to EST BE047599 (NID:g8364652) tz39c01.y1"
     misc_feature   594. .763
                    /note="match to EST BE908408 (NID:g10402954)"
     misc_feature   594. .763
                    /note="match to EST AA043371 (NID:g1521226) zk53e10.r1"
     misc_feature   594. .763
```

```
                     /note="match to EST AI670836 (NID:g4850567) wa04g10.x1"
     misc_feature    594. .763
                     /note="match to EST BF183086 (NID:g11061273)"
     misc_feature    594. .763
                     /note="similar to Homo sapiens EST BI114348
                     (NID:g14565249)"
     misc_feature    599. .763
                     /note="match to EST AL567345 (NID:g12920610)"
     misc_feature    617. .1084
                     /note="match to EST AA481361 (NID:g2210913) zv44e01.r1"
     misc_feature    622. .763
                     /note="match to EST AI860958 (NID:g5514574) wl56f05.x1"
     misc_feature    622. .763
                     /note="similar to Mus musculus EST AI196847 (NID:g3749453)
                     ui67e01.x1"
     misc_feature    684. .763
                     /note="match to EST BG992568 (NID:g14396638)"
     misc_feature    962. .1084
                     /note="match to EST AI656812 (NID:g4740791) tt54b06.x1"
     misc_feature    967. .1084
                     /note="match to EST BE908408 (NID:g10402954)"
     misc_feature    967. .1085
                     /note="match to EST BF304755 (NID:g11251653)"
     misc_feature    967. .1091
                     /note="match to EST AA043371 (NID:g1521226) zk53e10.r1"
     misc_feature    967. .1071

  Query Match             98.5%;  Score 1697;  DB 9;  Length 206854;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 1719;  Conservative    0;  Mismatches    0;  Indels    2;  Gaps    2;

Qy       1 cctggccccctcagatcaagaaggccttctttgctctggtggccaacggtgtgcgggcagc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  166697 CCTGGCCCCTCAGATCAAGAAGGCCTTCTTTGCTCTGGTGGCCAACGGTGTGCGGGCAGC 166638

Qy      61 ccctctatgggacagcaagaagcagagctttgtgggtgaggagaggctggggaggtgaag 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  166637 CCCTCTATGGGACAGCAAGAAGCAGAGCTTTGTGGGTGAGGAGAGGCTGGGGAGGTGAAG 166578

Qy     121 ggagatggaggaggtgaggggagatcttgtacggttgttctggggctgatctctgatat 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  166577 GGAGATGGAGGAGGTGAGGGGGAGATCTTGTACGGTTGTTCTGGGGCTGATCTCTGATAT 166518

Qy     181 accacaagcttggcttcaggccaagcccagccaggggccagggtggaggaaagtccatcc 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  166517 ACCACAAGCTTGGCTTCAGGCCAAGCCCAGCCAGGGGCCAGGGTGGAGGAAAGTCCATCC 166458

Qy     241 ggagtctgcatggccagctgggagaccctggggctcaatttccccatctgtggagccgct 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  166457 GGAGTCTGCATGGCCAGCTGGGAGACCCTGGGGCTCAATTTCCCCATCTGTGGAGCCGCT 166398

Qy     301 atgaccagctgacacctttcacctccgctactgcatggccctgtgccataggtgctaggg 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  166397 ATGACCAGCTGACACCTTTCACCTCCGCTACTGCATGGCCCTGTGCCATAGGTGCTAGGG 166338

Qy     361 agcaaatgggggaggcaggagagaaagagccccacttctcaggcctgggggctgcccc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  166337 AGCAAATGGGGGGAGGCAGGAGAGAAAGAGCCCCACTTCTCAGGCCTGGGGGGCTGCCCC 166278

Qy     421 actgtcctgttcccacagtccccactgtgtctcagcacaaggacactggcagggtgggga 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  166277 ACTGTCCTGTTCCCACAGTCCCCACTGTGTCTCAGCACAAGGACACTGGCAGGGTGGGGA 166218

Qy     481 ggggatctgaccctcaacctgccttccacccaaaggccccgggctgacctcctccccgcc 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  166217 GGGGATCTGACCCTCAACCTGCCTTCCACCCAAAGGCCCCGGGCTGACCTCCTCCCCGCC 166158

Qy     541 cctcccctgcagggatgctgaccatcactgacttcatcctggtgctgcatcgctactaca 600
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  166157 CCT-CCCTGCAGGGATGCTGACCATCACTGACTTCATCCTGGTGCTGCATCGCTACTACA 166099

Qy     601 ggtcccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcgg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db  166098 GGTCCCCCCTGGTGAGGAGTGGGCTGGGAATCTTATGGGCACCCAGAGGGGCGGGGGCGG 166039

Qy     661 aggggagtcctcctggagcctggtgccctagaagcccacgtctttctgacttctggagtc 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  166038 AGGGGAGTCCTCCTGGAGCCTGGTGCCCTAGAAGCCCACGTCTTTCTGACTTCTGGAGTC 165979

Qy     721 ctgtcgatgtctctaggtccagatctatgagattgaacaacataagattgagacctggag 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  165978 CTGTCGATGTCTCTAGGTCCAGATCTATGAGATTGAACAACATAAGATTGAGACCTGGAG 165919

Qy     781 gggtgagtggggagaggaacccggaaaaggggctgttggtgatggtgggccagggcttaag 840
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||||||
Db  165918 GGGTGAGTGGGGAGAGGAACCCGGAAA-GGGCTGTTGGTGATGGTGGGCCAGGGCTTAAG 165860

Qy     841 gtggaggatgggcagtggggatgtcctggagtgaacaggggagggacaataggagcctcg 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  165859 GTGGAGGATGGGCAGTGGGGATGTCCTGGAGTGAACAGGGGAGGGACAATAGGAGCCTCG 165800

Qy     901 ggtgcctgacggaagggaagctgcctgggactgcaaggtgaggcaggtgaccggctcccc 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  165799 GGTGCCTGACGGAAGGGAAGCTGCCTGGGACTGCAAGGTGAGGCAGGTGACCGGCTCCCC 165740

Qy     961 tggcctgactctggctctttctgcagagatctacctgcaaggctgcttcaagcctctggt 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  165739 TGGCCTGACTCTGGCTCTTTCTGCAGAGATCTACCTGCAAGGCTGCTTCAAGCCTCTGGT 165680

Qy    1021 ctccatctctcctaatgataggtgggtgtctctgctcattcacctgagcctcctcctccc 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  165679 CTCCATCTCTCCTAATGATAGGTGGGTGTCTCTGCTCATTCACCTGAGCCTCCTCCTCCC 165620

Qy    1081 acagtccccttccccagtcccactcagctctgaactcacctcttcatcctaggcggcaca 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  165619 ACAGTCCCCTTCCCCAGTCCCACTCAGCTCTGAACTCACCTCTTCATCCTAGGCGGCACA 165560

Qy    1141 cagacaagggagccttggtgccctgccctccttttttaggggcctgggatggaggttgtct 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  165559 CAGACAAGGGAGCCTTGGTGCCCTGCCCTCCTTTTTAGGGGCCTGGGATGGAGGTTGTCT 165500

Qy    1201 ctccctaggctgccccgaggctcactgctcccatctctgcagcctgtttgaagctgtcta 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  165499 CTCCCTAGGCTGCCCCGAGGCTCACTGCTCCCATCTCTGCAGCCTGTTTGAAGCTGTCTA 165440

Qy    1261 caccctcatcaagaaccggatccatcgcctgcctgttcttgacccggtgtcaggcaacgt 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  165439 CACCCTCATCAAGAACCGGATCCATCGCCTGCCTGTTCTTGACCCGGTGTCAGGCAACGT 165380

Qy    1321 actccacatcctcacacacaaacgcctgctcaagttcctgcacatctttgtaagcctggg 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  165379 ACTCCACATCCTCACACACAAACGCCTGCTCAAGTTCCTGCACATCTTTGTAAGCCTGGG 165320

Qy    1381 cccaggtgggaggaaggggagacctgggcaggtgatcagagggcctgaggagtcttcag 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  165319 CCCAGGTGGGAGGAAGGGGGAGACCTGGGCAGGTGATCAGAGGGCCTGAGGAGTCTTCAG 165260

Qy    1441 ccctagcagtcgtggggaagagctgggagccctcttgaagctgctggatccctgatctcc 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  165259 CCCTAGCAGTCGTGGGGAAGAGCTGGGAGCCCTCTTGAAGCTGCTGGATCCCTGATCTCC 165200

Qy    1501 acctggtccccatcctaaccagggttccctgctgccccggcccctccttcctctaccgcac 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  165199 ACCTGGTCCCCATCCTAACCAGGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCAC 165140

Qy    1561 tatccaagatttgggcatcggcacattccgagacttggctgtggtgctggagacagcacc 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  165139 TATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCTGTGGTGCTGGAGACAGCACC 165080

Qy    1621 catcctgactgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaa 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  165079 CATCCTGACTGCACTGGACATCTTTGTGGACCGGCGTGTGTCTGCACTGCCTGTGGTCAA 165020

Qy    1681 cgaatgtggtacccacccccaggatgagaggctcgggctgg 1721
          ||||||||||||||||||||||||||||||||||||||||
Db  165019 CGAATGTGGTACCCACCCCCAGGATGAGAGGCTCGGGCTGG 164979
```

```
RESULT   5
AF214521
LOCUS       AF214521                5888 bp    DNA     linear   MAM 03-JUN-2000
DEFINITION  Sus scrofa AMPK gamma subunit (PRKAG3) gene, complete cds.
ACCESSION   AF214521
VERSION     AF214521.1  GI:8215685
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 5888)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A., Thelander,M.,
            Rogel-Gaillard,C., Paul,S., Iannuccelli,N., Rask,L., Ronne,H.,
            Lundstrom,K., Reinsch,N., Gellin,J., Kalm,E., Roy,P.L., Chardon,P.
            and Andersson,L.
  TITLE     A mutation in PRKAG3 associated with excess glycogen content in pig
            skeletal muscle
  JOURNAL   Science 288 (5469), 1248-1251 (2000)
  MEDLINE   20280150
   PUBMED   10818001
REFERENCE   2  (bases 1 to 5888)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A.,
            Rogel-Gaillard,C., Paul,S., Gellin,J., Lundstrom,K., Reinsch,N.,
            Kalm,E., Le Roy,P., Chardon,P. and Andersson,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-1999) Dept Animal Breeding and Genetics, Swedish
            University of Agricultural Sciences, BMC Box 597, Uppsala 751 24,
            Sweden
FEATURES             Location/Qualifiers
     source          1. .5888
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /chromosome="15"
                     /map="15q"
     mRNA            join(<1. .154,515. .918,1809. .1890,2349. .2407,2509. .2554,
                     2771. .2825,3027. .3153,3286. .3451,4578. .4615,4791. .4937,
                     5294. .>5410)
                     /gene="PRKAG3"
                     /product="AMPK gamma subunit"
     gene            <1. .>5410
                     /gene="PRKAG3"
     CDS             join(1. .154,515. .918,1809. .1890,2349. .2407,2509. .2554,
                     2771. .2825,3027. .3153,3286. .3451,4578. .4615,4791. .4937,
                     5294. .5410)
                     /gene="PRKAG3"
                     /codon_start=1
                     /product="AMPK gamma subunit"
                     /protein_id="AAF73989.1"
                     /db_xref="GI:8215686"
                     /translation="MSFLEQGESRSWPSRAVTTSSERSHGDQGNKASRWTRQEDVEEG
                     GPPGPREGPQSRPVAESTGQEATFPKATPLAQAAPLAEVDNPPTERDILPSDCAASAS
                     DSNTDHLDLGIEFSASAASGDELGLVEEKPAPCPSPEVLLPRLGWDDELQKPGAQVYM
                     HFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEEHKIETWREIYLQGCFKPLVSISPNDSLFEAVYALI
                     KNRIHRLPVLDPVSGAVLHILTHKRLLKFLHIFGTLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNETGQVVGLYSRFDVIHLAAQQTYNHLDMNV
                     GEALRQRTLCLEGVLSCQPHETLGEVIDRIVREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT     1274 a   1634 c   1638 g   1339 t      3 others
ORIGIN
```

```
  Query Match             35.6%;  Score 612.6;  DB 4;  Length 5888;
  Best Local Similarity   66.3%;  Pred. No. 1e-137;
  Matches 1155;  Conservative    0;  Mismatches  484;  Indels  102;  Gaps   15;

Qy       5 gccccctcagatcaagaaggccttctttgctctggtggccaacggtgtgcgggcagcccct 64
           | | ||||||||||||||||||||||||||||| |||||||||||||||| || || || || |||
Db    1800 GGCGCTCAGATCAAGAAGGCCTTCTTTGCCCTGGTGGCCAACGGCGTCCGAGCGGCACCT 1859

Qy      65 ctatgggacagcaagaagcagagctttgtgggtgaggagaggctggggaggtgaagggag 124
           | |||||||||||||||||||||||||| |||||||||||| ||||||||||  |  || |
Db    1860 TTGTGGGACAGCAAGAAGCAGAGCTTCGTGGGTGAGGAGGGGCTGGGGAGG-CAGAGGTG 1918

Qy     125 atggaggaggtgaggggggagatcttgtacggttgttctggggctgatctctgatatacca 184
           ||| | |||  |  ||| || |||||  |||  |||| |||| || |||||| | ||||
Db    1919 GTGGGGAAGGGAATAGGGGGACCTTGTGGGGTGATTCTAGGGCCGAGCTCTGACACACCA 1978

Qy     185 caagcttggcttcaggccaagcccagccagggggccagggtggaggaaagtccatccggag 244
           || ||||      ||     |||  ||| || |  |||| ||  | |  |    ||||
Db    1979 CAGGCTTCAACCAAGCAGGGGCCTGGCCTGGAG--AGGGGGGGAGCATTTGACCCCGGTC 2036

Qy     245 tctgcatggccagctgggagaccctggggctcaatttccccatctgtggagccgctatga 304
           ||    |||||||||||||||                  | || |||| |    ||| |||
Db    2037 TCCTGGTGGCCAGCTGGGAGA---------------TCTCAACTGTAGGAGAGCTGTGA 2080

Qy     305 ccagctgacacctttcacctccgctactgcatggccctgtgcca------taggtgctag 358
           |||||||| |||    |        |  | | ||||||  ||       ||     ||
Db    2081 CCAGCTGACCCCTCCAGCTCTACTACCCCAAGGTCCCTGTCGCAGGTGCTAAGTAAGAAG 2140

Qy     359 ggagcaaatgggggaggcaggagagaaagagccccacttctcaggcctgggggctgcc 418
           |  ||   || || ||| ||     ||| ||    |        |   || | |
Db    2141 AGGACAGGCGGAGGAAGGAAGTCAGAAAATAGAAGAAGCAGGGCAGGAAGGAGAGAAATG 2200

Qy     419 ccactgtcctgttcccacagtccccactgtgtctcagcacaaggacactggcagggtggg 478
           ||  |      |  | | |  | | | |||     |||          |
Db    2201 ACAGGGGAAGCATAAGAGGGACAACCCCATTTGTCAGGCACGGGAGGGGCTGCCCTCCTG 2260

Qy     479 gaggggatctgaccctcaacctgcc------------ttccacccaaaggccccgggct 525
                | | | | |    |                          | || ||||||||
Db    2261 TCCTCTTTTGGCCACCCTCAGTAAAAGGATGTGGGCAGGGTGGGGGGAGGGGCCCGGGCT 2320

Qy     526 gacctcctccccgcccctc-ccctgcagggatgctgaccatcactgacttcatcctggtg 584
           |||| ||    | |||||| |||  |||||||||||||||||| |||||||| |||||
Db    2321 GACCCCCATTGCTCCCCTCGCCCCACAGGGATGCTGACCATCACAGACTTCATCTTGGTG 2380

Qy     585 ctgcatcgctactacaggtccccccctggtgaggagtgggctgggaatcttatgggcaccc 644
           ||||| ||||| |||||||||||||||||||||||||||| |||||  || |  |  |||||
Db    2381 CTGCACCGCTATTACAGGTCCCCCCTGGTGAGGAGTGGTCTGGGGGTCCT-GGAACACCC 2439

Qy     645 agaggggcgggggcggaggggagtcctcctggagcctggtgccctagaagcccacgtctt 704
           |   |||| |||| ||| ||   ||                          ||   ||
Db    2440 ATCTGGGCTGGGGTGGAAGGAGTTC----------------------AGGGGACCCTCG 2476

Qy     705 tctgacttctggagtcctgtcgatgtctctaggtccagatctatgagattgaacaacata 764
           |||||||  ||||| | || | ||||| ||||||||||||||| |||||||| ||||||
Db    2477 CCTGACTTTGGGAGTTCCGTTGCTGTCTTTAGGTCCAGATCTACGAGATTGAAGAACATA 2536

Qy     765 agattgagacctggaggggtgagtggggagaggaacccggaaaggggctgttggt----- 819
           ||||||||||||||||||||||  ||    || ||  |  |||||||||  |||
Db    2537 AGATTGAGACCTGGAGGGGTGAGCAGGCGAGGGGACGGGCGAAGGGGCTGAGGGTGTGTG 2596

Qy     820 -----gatggtgggccagggcttaaggtggaggatgggcagtggggatgtcctggagtga 874
                |   | || | ||| | | |||  ||| ||| ||||||| | | ||||||||  |
Db    2597 GGTGAGGATGGGGCCAAGGACCTCAGGGAGAGCATGCGCAGTGGAGGTTTCCTGGAGGAA 2656

Qy     875 acaggggagggacaataggagcctcgggtgcctgacggaagggaagctgcctgggactg- 933
           | || |  ||   || ||   | | || |    | || |||||  | | |||||  ||
Db    2657 GCGGGAGGAGGGTGATCGGGAGCCCAGGGGATCTAAGGGAGGGAGACAGTCTGGGGGTGG 2716

Qy     934 -caaggtgaggcaggtgaccggctcccctggcctgactctggctctttctgcagagatct 992
           || |    |  |||| |||| ||  ||  |||| ||||||| || ||||||||||||
Db    2717 CCACGTGAGGCGGGGTGGTCGGCCCCTTTGTGCTGATTCTGGCTTTTCCTGCAGAGATCT 2776

Qy     993 acctgcaaggctgcttcaagcctctggtctccatctctcctaatgataggtgggtgtctc 1052
           |||| |||||||||||||||||||||||||||||||||||| ||||| ||||| |  || |
Db    2777 ACCTTCAAGGCTGCTTCAAGCCTCTGGTCTCCATCTCTCCCAATGACAGGTGAGCTTCCC 2836

Qy    1053 tgctcattcacctgagcctcctcctcccacagtccccttccccagtcccactcagctctg 1112
              |  |||  |||||||||   ||| ||    ||    |     |   || | |||
Db    2837 CAGCCGCCCACTCGAGCCTCCTTGCCCCGCACAGACCCCTTCTCCAGCTCATCGGTTCTA 2896

Qy    1113 aactcacctcttcatcctaggcggcacacagacaagggagccttggtgccctgccctcct 1172
```

```
Db    2897 AGCTCATGGACTCATCGTCCGTGGACTGCAGATGCGGCAGCTTTGACACCCTGTCCTCCT 2956

Qy    1173 ttttag-gggcctgggatqgaggttgtctctccctaqgctgccccgaggctcactgctcc 1231
Db    2957 CTCCAGGGGGGCTGGGATGAAGGGGCTCTCTTTCCAGACTGCCCC-AGGCTCACTGCTCC 3015

Qy    1232 catctctgcagcctgtttgaagctgtctacaccctcatcaagaaccggatccatcgcctg 1291
Db    3016 CACCTCCACAGCCTGTTCGAAGCTGTCTACGCCCTCATCAAGAACCGGATCCACCGCCTG 3075

Qy    1292 cctgttcttgacccggtgtcaggcaacgtactccacatcctcacacacaaacgcctgctc 1351
Db    3076 CCGGTCCTGGACCCTGTCTCCGGGGCTGTGCTCCACATCCTCACACATAAGCGGCTTCTC 3135

Qy    1352 aagttcctgcacatctttgtaagcctgggcccaggtgggaggaaggggagacctgggca 1411
Db    3136 AAGTTCCTGCACATCTTTGTGAGCCTGGGCACAGCCTCAGGGA---------------- 3178

Qy    1412 ggtgatcagagggcctgaggagtcttcagccctagcagtcgtg--gggaagagctgggag 1469
Db    3179 --CAACCTGAGTGGCTGAGAAGTCTTCAGCCCTAGGGATGGGGGAGGGAGTAGCTGGGAG 3236

Qy    1470 ccctcttgaagctgctggatccctgatctccacctggtcccatcctaaccagggttccc 1529
Db    3237 CCCCCTGAGGGCTAC----TCCCTCCTGGCCTCACCTGTCCCAACCCAACCAGGGCACCC 3292

Qy    1530 tqctgccccggccctccttcctctaccgcactatccaagatttgggcatcggcacattcc 1589
Db    3293 TGCTGCCCCGGCCCTCCTTCCTCTACCGCACCATCCAAGATTTGGGCATCGGCACATTCC 3352

Qy    1590 gagacttggctgtggtgctggagacagcacccatcctgactgcactggacatctttgtgg 1649
Db    3353 GAGACTTGGCCGTGGTGCTGGAAACGGCGCCCATCCTGACCGCACTGGACATCTTCGTGG 3412

Qy    1650 accggcgtgtgtctgcactgcctgtggtcaacgaatgtggtacccacccccaggatgaga 1709
Db    3413 ACCGGCGTGTGTCTGCGCTGCCTGTGGTCAACGAAACTGGTACCTATGCCCAGGATGGGG 3472

Qy    1710 q 1710
Db    3473 G 3473


RESULT   6
AF336381/c
LOCUS       AF336381               227724 bp    DNA     linear   HTG 02-APR-2001
DEFINITION  Mus musculus chromosome 1 clone PAC510; PAC457, *** SEQUENCING IN
            PROGRESS ***, 3 unordered pieces.
ACCESSION   AF336381
VERSION     AF336381.1  GI:13507298
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 227724)
  AUTHORS   Rump,A., Hayes,C., Brown,S.D.M. and Rosenthal,A.
  TITLE     Mouse chromosome 1 genomic sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 227724)
  AUTHORS   Rump,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JAN-2001) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenberstr. 11, Jena 07745, Germany
COMMENT     * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    17869: contig of 17869 bp in length
            *    17870    17969: gap of unknown length
            *    17970    32746: contig of 14777 bp in length
            *    32747    32846: gap of unknown length
            *    32847   227724: contig of 194878 bp in length.
FEATURES             Location/Qualifiers
     source          1. .227724
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="1"
                     /clone="PAC510; PAC457"
BASE COUNT    57663 a  55423 c  56238 g  58199 t    201 others
ORIGIN


  Query Match           12.2%;  Score 209.4;  DB 2;  Length 227724;
  Best Local Similarity   67.6%;  Pred. No. 2.3e-40;
  Matches  401;  Conservative    0;  Mismatches  126;  Indels   66;  Gaps    5;

Qy    1100 ccactcagctctgaactcacctcttcatcctaggcggcacacagacaagggagccttggt 1159
Db    5890 CCTCTCGGTCCTGCATTCACTCACTCCTTCCGGGAAGTAGAGGGATGAGGGAACCTTGGT 5831

Qy    1160 gccctgccctcctttttaggggcctgggatggaggttgtctctccctaggctgccccgag 1219
Db    5830 TCCTTG--------TTAGGGGGCCTCAGATGGAGGGTTTCT---ACCACGGTGCTCCTAG 5782

Qy    1220 gctcactgctccca-tctctgcagcctgtttgaagctgtctacaccctcatcaagaaccg 1278
Db    5781 TCTCACTGTTTCCATTCCCTAAAGCCTGTTTGAAGCTGTCTATGCCCTCATCAAGAACCG 5722

Qy    1279 gatccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctcacaca 1338
Db    5721 AATCCACCGCCTGCCGGTCCTGGACCCGGTCTCTGGCACTGTGCTCTACATACTCACACA 5662

Qy    1339 caaacgcctgctcaagttcctgcacatctttgtaagcctgggcccaggtgggaggaaggg 1398
Db    5661 CAAGCGGCTACTCAAGTTCCTGCATATATTTGTGAGCCTGGG--ATGGTGGGAGCAAGGG 5604

Qy    1399 ggagacctgggcaggtgatcagagggcctgaggagtcttcagccctagcagtcgtgggga 1458
Db    5603 GAGAGCCACGGGAGACACT---------------------------------------- 5583

Qy    1459 agagctgggagccctcttgaagctgctggatccctgatctccacctggtccccatcctaa 1518
Db    5584 ----------CTAGATCAAGTGTCTTGATCTGATGAGATCACTACCATGTCCCATCCTA 5536

Qy    1519 ccagggttccctgctgccccggccctccttcctctaccgcactatccaagatttgggcat 1578
Db    5535 ACAGGGTGCCCTGTTGCCCCGGCCCTCCTTCCTCTGCCGCACTATCCAAGACTTGGGCAT 5476

Qy    1579 cggcacattccgagacttggctgtggtgctggagacagcacccatcctgactgcactgga 1638
Db    5475 CGGCACATTCCGAGATTTGGCTGTAGTTCTGGAAACAGCTCCTGTCCTGACTGCGCTGGA 5416

Qy    1639 catctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtggta 1691
Db    5415 CATCTTTGTGGACCGACGTGTGTCTGCACTGCCTGTGGTCAATGAATCTGGTA 5363


RESULT   7
AX281582
LOCUS       AX281582               1647 bp    DNA     linear   PAT 02-NOV-2001
DEFINITION  Sequence 5 from Patent WO0177305.
ACCESSION   AX281582
VERSION     AX281582.1  GI:16608833
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Andersson,L., Luthman,H. and Marklund,S.
```

```
  TITLE     Variants of the human amp-activated protein kinase gamma 3 subunit
  JOURNAL   Patent: WO 0177305-A 5 18-OCT-2001;
            Arexis AB (SE)
FEATURES             Location/Qualifiers
     source          1. .1647
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             20. .1489
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD10589.1"
                     /db_xref="GI:16608834"
                     /translation="MEPGLEHALRRTPSWSSLGGSEHQEMSFLEQENSSSWPSPAVTS
                     SSERIRGKRRAKALRWTRQKSVEEGEPPGQGEGPRSRPAAESTGLEATFPKTTPLAQA
                     DPAGVGTPPTGWDCLPSDCTASAAGSSTDDVELATEFPATEAWECELEGLLEERPALC
                     LSPQAPFPKLGWDDELRKPGAQIYMRFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFA
                     LVANGVRAAPLWDSKKQSFVGMLTITDFILVLHRYYRSPLVQIYEIEQHKIETWREIY
                     LQGCFKPLVSISPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIF
                     GSLLPRPSFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSALPVVNECGQV
                     VGLYSRFDVIHLAAQQTYNHLDMSVGEALRQRTLCLEGVLSCQPHESLGEVIDRIARE
                     QVHRLVLVDETQHLLGVVSLSDILQALVLSPAGIDALGA"
BASE COUNT      346 a    502 c    462 g    337 t
ORIGIN


  Query Match             9.8%;  Score 168.2;  DB 6;  Length 1647;
  Best Local Similarity   95.6%;  Pred. No. 4.2e-30;
  Matches  173;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    1510 ccatcctaaccagggttccctgctgccccggccctccttcctctaccgcactatccaaaga 1569
           ||  |  | |   ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1009 CCTGCACATCTTTGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACTATCCAAGA 1068

Qy    1570 tttgggcatcggcacattccgagacttggctgtggtgctggagacagcacccatcctgac 1629
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1069 TTTGGGCATCGGCACATTCCGAGACTTGGCTGTGGTGCTGGAGACAGCACCCATCCTGAC 1128

Qy    1630 tgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtgg 1689
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1129 TGCACTGGACATCTTTGTGGACCGGCGTGTGTCTGCACTGCCTGTGGTCAACGAATGTGG 1188

Qy    1690 t 1690
           |
Db    1189 T 1189


RESULT   8
AX099776
LOCUS       AX099776                2109 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 3 from Patent WO0120003.
ACCESSION   AX099776
VERSION     AX099776.1  GI:13538810
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2109)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 3 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .2109
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             472. .1389
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35799.1"
                     /db_xref="GI:13538811"
                     /translation="MRFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAP
                     LWDSKKQSFVGMLTITDFILVLHRYYRSPLVQIYEIEQHKIETWREIYLQGCFKPLVS
                     ISPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLPRPSFL
                     YRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSALPVVNECGQVVGLYSRFDVI
                     HLAAQQTYNHLDMSVGEALRQRTLCLEGVLSCQPHESLGEVIDRIAREQVHRLVLVDE
                     TQHLLGVVSLSDILQALVLSPAGIDALGA"
BASE COUNT      458 a    621 c    560 g    470 t
ORIGIN


  Query Match             9.8%;  Score 168.2;  DB 6;  Length 2109;
  Best Local Similarity   95.6%;  Pred. No. 4e-30;
  Matches  173;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    1510 ccatcctaaccagggttccctgctgccccggccctccttcctctaccgcactatccaaaga 1569
           ||  |  | |   ||||||||||||||||||||||||||||||||||||||||||||||||
Db     909 CCTGCACATCTTTGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACTATCCAAGA 968

Qy    1570 tttgggcatcggcacattccgagacttggctgtggtgctggagacagcacccatcctgac 1629
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     969 TTTGGGCATCGGCACATTCCGAGACTTGGCTGTGGTGCTGGAGACAGCACCCATCCTGAC 1028

Qy    1630 tgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtgg 1689
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1029 TGCACTGGACATCTTTGTGGACCGGCGTGTGTCTGCACTGCCTGTGGTCAACGAATGTGG 1088

Qy    1690 t 1690
           |
Db    1089 T 1089


RESULT   9
AX099802
LOCUS       AX099802                2115 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 29 from Patent WO0120003.
ACCESSION   AX099802
VERSION     AX099802.1  GI:13538836
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2115)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 29 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .2115
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             1. .1395
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35801.1"
                     /db_xref="GI:13538837"
                     /translation="MSFLEQENSSSWPSPAVTSSSERIRGKRRAKALRWTRQKSVEEG
                     EPPGQGEGPRSRPTAESTGLEATFPKTTPLAQADPAGVGTPPTGWDCLPSDCTASAAG
                     SSTDDVELATEFPATEAWECELEGLLEERPALCLSPQAPFPKLGWDDELRKPGAQIYM
                     RFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEQHKIETWREIYLQGCFKPLVSISPNDSLFEAVYTLI
                     KNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNECGQVVGLYSRFDVIHLAAQQTYNHLDMSV
                     GEALRQRTLCLEGVLSCQPHESLGEVIDRIAREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT      460 a    622 c    562 g    471 t
ORIGIN
```

```
  Query Match             9.8%;  Score 168.2;  DB 6;  Length 2115;
  Best Local Similarity   95.6%;  Pred. No. 4e-30;
  Matches  173;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   1510 ccatcctaaccagggttccctgctgccccggccctccttcctctaccgcactatccaaga 1569
          ||  |  | |   ||||||||||||||||||||||||||||||||||||||||||||||||
Db    915 CCTGCACATCTTTGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACTATCCAAGA 974

Qy   1570 tttgggcatcggcacattccgagacttggctgtggtgctggagacagcacccatcctgac 1629
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    975 TTTGGGCATCGGCACATTCCGAGACTTGGCTGTGGTGCTGGAGACAGCACCCATCCTGAC 1034

Qy   1630 tgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtgg 1689
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1035 TGCACTGGACATCTTTGTGGACCGGCGTGTGTCTGCACTGCCTGTGGTCAACGAATGTGG 1094

Qy   1690 t 1690
          |
Db   1095 T 1095


RESULT  10
AF214519
LOCUS       AF214519                2115 bp    mRNA    linear   PRI 03-JUN-2000
DEFINITION  Homo sapiens AMP-activated protein kinase gamma subunit (PRKAG3)
            mRNA, complete cds.
ACCESSION   AF214519
VERSION     AF214519.1  GI:8215681
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2115)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A., Thelander,M.,
            Rogel-Gaillard,C., Paul,S., Iannuccelli,N., Rask,L., Ronne,H.,
            Lundstrom,K., Reinsch,N., Gellin,J., Kalm,E., Roy,P.L., Chardon,P.
            and Andersson,L.
  TITLE     A mutation in PRKAG3 associated with excess glycogen content in pig
            skeletal muscle
  JOURNAL   Science 288 (5469), 1248-1251 (2000)
  MEDLINE   20280150
   PUBMED   10818001
REFERENCE   2  (bases 1 to 2115)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A.,
            Rogel-Gaillard,C., Paul,S., Gellin,J., Lundstrom,K., Reinsch,N.,
            Kalm,E., Le Roy,P., Chardon,P. and Andersson,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-1999) Dept Animal Breeding and Genetics, Swedish
            University of Agricultural Sciences, BMC box 597, Uppsala 751 24,
            Sweden
FEATURES             Location/Qualifiers
     source          1. .2115
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2p"
                     /tissue_type="skeletal muscle"
     gene            1. .2115
                     /gene="PRKAG3"
     CDS             1. .1395
                     /gene="PRKAG3"
                     /note="AMPKG3"
                     /codon_start=1
                     /product="AMP-activated protein kinase gamma subunit"
                     /protein_id="AAF73987.1"
                     /db_xref="GI:8215682"
                     /translation="MSFLEQENSSSWPSPAVTSSSERIRGKRRAKALRWTRQKSVEEG
                     EPPGQGEGPRSRPTAESTGLEATFPKTTPLAQADPAGVGTPPTGWDCLPSDCTASAAG
                     SSTDDVELATEFPATEAWECELEGLLEERPALCLSPQAPFPKLGWDDELRKPGAQIYM
                     RFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEQHKIETWREIYLQGCFKPLVSISPNDSLFEAVYTLI
                     KNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNECGQVVGLYSRFDVIHLAAQQTYNHLDMSV
                     GEALRQRTLCLEGVLSCQPHESLGEVIDRIAREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT      460 a    622 c    562 g    471 t
ORIGIN


  Query Match             9.8%;  Score 168.2;  DB 9;  Length 2115;
  Best Local Similarity   95.6%;  Pred. No. 4e-30;
  Matches  173;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   1510 ccatcctaaccagggttccctgctgccccggccctccttcctctaccgcactatccaaga 1569
          ||  |  | |   ||||||||||||||||||||||||||||||||||||||||||||||||
Db    915 CCTGCACATCTTTGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACTATCCAAGA 974

Qy   1570 tttgggcatcggcacattccgagacttggctgtggtgctggagacagcacccatcctgac 1629
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    975 TTTGGGCATCGGCACATTCCGAGACTTGGCTGTGGTGCTGGAGACAGCACCCATCCTGAC 1034

Qy   1630 tgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtgg 1689
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1035 TGCACTGGACATCTTTGTGGACCGGCGTGTGTCTGCACTGCCTGTGGTCAACGAATGTGG 1094

Qy   1690 t 1690
          |
Db   1095 T 1095


RESULT  11
HSA249977
LOCUS       HSA249977               2290 bp    mRNA    linear   PRI 07-APR-2000
DEFINITION  Homo sapiens mRNA for AMP-activated protein kinase gamma 3 subunit
            (AMPK gamma 3 gene).
ACCESSION   AJ249977
VERSION     AJ249977.1  GI:6688200
KEYWORDS    AMP-activated protein kinase; AMPK gamma 3 gene; gamma 3 subunit.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2290)
  AUTHORS   Cheung,P.C., Salt,I.P., Davies,S.P., Hardie,D.G. and Carling,D.
  TITLE     Characterization of AMP-activated protein kinase gamma-subunit
            isoforms and their role in AMP binding
  JOURNAL   Biochem. J. 346 Pt 3, 659-669 (2000)
  MEDLINE   20164049
REFERENCE   2  (bases 1 to 2290)
  AUTHORS   Carling,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-OCT-1999) Carling D., Cellular Stress Group, MRC
            Clinical Sciences Centre, Hammersmith Hospital, DuCane Road,
            London, W12 0NN, UNITED KINGDOM
FEATURES             Location/Qualifiers
     source          1. .2290
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            22. .1500
                     /gene="AMPK gamma 3"
     CDS             22. .1500
                     /gene="AMPK gamma 3"
                     /function="AMP-activated protein kinase regulatory
                     subunit"
                     /codon_start=1
                     /evidence=experimental
                     /product="AMP-activated protein kinase gamma 3 subunit"
                     /protein_id="CAB65117.1"
                     /db_xref="GI:6688201"
                     /translation="MEPGLEHALRRTPSWSSLGGSEHQEMSFLEQENSSSWPSPAVTS
                     SSERIRGKRRAKALRWTRQKSVEEGEPPGQGEGPRSRPAAESTGLEATFPKTTPLAQA
                     DPAGVGTPPTGWDCLPSDCTASAAGSSTDDVELATEFPATEAWECELEGLLEERPALC
                     LSPQAPFPKLGWDDELRKPGAQIYMRFIEEHTCYDAMATSSKLVIFDTMLEIKKAFFA
```

```
                    LVANGVRAAPLWDSKKQSFVGMLTITDFILVLHRYYRSPLVQIYEIEQHKIETWREIY
                    LQGCFKPLVSISPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIF
                    GSLLPRPSFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSALPVVNECGQV
                    VGLYSRFDVIHLAAQQTYNHLDMSVGEALRKRTLCLEGVLSCQPHESLGEVIDRIARE
                    QVHRLVLVDETQHLLGVVSLSDILQALVLSPAGIDPSGPEKI"
BASE COUNT      501 a    674 c    617 g    498 t
ORIGIN


  Query Match             9.8%;  Score 168.2;  DB 9;  Length 2290;
  Best Local Similarity   95.6%;  Pred. No. 4e-30;
  Matches  173;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    1510 ccatcctaaccagggttccctgctgccccggccctccttcctctaccgcactatccaaga 1569
           ||  |  | |   ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1011 CCTGCACATCTTTGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACTATCCAAGA 1070

Qy    1570 tttgggcatcggcacattccgagacttggctgtggtgctggagacagcacccatcctgac 1629
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1071 TTTGGGCATCGGCACATTCCGAGACTTGGCTGTGGTGCTGGAGACAGCACCCATCCTGAC 1130

Qy    1630 tgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtgg 1689
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1131 TGCACTGGACATCTTTGTGGACCGGCGTGTGTCTGCACTGCCTGTGGTCAACGAATGTGG 1190

Qy    1690 t 1690
           |
Db    1191 T 1191


RESULT  12
AX099774
LOCUS       AX099774               1867 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 1 from Patent WO0120003.
ACCESSION   AX099774
VERSION     AX099774.1  GI:13538808
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 1867)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 1 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .1867
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
     CDS             472. .1389
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35798.1"
                     /db_xref="GI:13538809"
                     /translation="MHFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAP
                     LWDSKKQSFVGMLTITDFILVLHRYYRSPLVQIYEIEEHKIETWREIYLQGCFKPLVS
                     ISPNDSLFEAVYALIKNRIHRLPVLDPVSGAVLHILTHKRLLKFLHIFGTLLPRPSFL
                     YRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSALPVVNETGQVVGLYSRFDVI
                     HLAAQQTYNHLDMNVGEALRQRTLCLEGVLSCQPHETLGEVIDRIVREQVHRLVLVDE
                     TQHLLGVVSLSDILQALVLSPAGIDALGA"
BASE COUNT      380 a    583 c    529 g    375 t
ORIGIN


  Query Match             8.6%;  Score 148;  DB 6;  Length 1867;
  Best Local Similarity   88.9%;  Pred. No. 3.2e-25;
  Matches  160;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy    1510 ccatcctaaccagggttccctgctgccccggccctccttcctctaccgcactatccaaga 1569
           ||  |  | |   ||  ||||||||||||||||||||||||||||||||||| ||||||||
Db     909 CCTGCACATCTTTGGCACCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACCATCCAAGA 968

Qy    1570 tttgggcatcggcacattccgagacttggctgtggtgctggagacagcacccatcctgac 1629
           ||||||||||||||||||||||||||||||| |||||||||| || || ||||||||||
Db     969 TTTGGGCATCGGCACATTCCGAGACTTGGCCGTGGTGCTGGAAACGGCGCCCATCCTGAC 1028

Qy    1630 tgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtgg 1689
           |||||||||||||| |||||||||||||||||||| |||||||||||||||||  |||
Db    1029 CGCACTGGACATCTTCGTGGACCGGCGTGTGTCTGCGCTGCCTGTGGTCAACGAAACTGG 1088


RESULT  13
AF214520
LOCUS       AF214520               1873 bp    mRNA    linear   MAM 03-JUN-2000
DEFINITION  Sus scrofa AMP-activated protein kinase gamma subunit (PRKAG3)
            mRNA, complete cds.
ACCESSION   AF214520
VERSION     AF214520.1  GI:8215683
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 1873)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A., Thelander,M.,
            Rogel-Gaillard,C., Paul,S., Iannuccelli,N., Rask,L., Ronne,H.,
            Lundstrom,K., Reinsch,N., Gellin,J., Kalm,E., Roy,P.L., Chardon,P.
            and Andersson,L.
  TITLE     A mutation in PRKAG3 associated with excess glycogen content in pig
            skeletal muscle
  JOURNAL   Science 288 (5469), 1248-1251 (2000)
  MEDLINE   20280150
   PUBMED   10818001
REFERENCE   2  (bases 1 to 1873)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A.,
            Rogel-Gaillard,C., Paul,S., Gellin,J., Lundstrom,K., Reinsch,N.,
            Kalm,E., Le Roy,P., Chardon,P. and Andersson,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-1999) Dept Animal Breeding and Genetics, Swedish
            University of Agricultural Sciences, BMC box 597, Uppsala 751 24,
            Sweden
FEATURES             Location/Qualifiers
     source          1. .1873
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /chromosome="15"
                     /map="15q"
                     /tissue_type="skeletal muscle"
     gene            1. .1873
                     /gene="PRKAG3"
     CDS             1. .1395
                     /gene="PRKAG3"
                     /note="AMPKG3"
                     /codon_start=1
                     /product="AMP-activated protein kinase gamma subunit"
                     /protein_id="AAF73988.1"
                     /db_xref="GI:8215684"
                     /translation="MSFLEQGESRSWPSRAVTTSSERSHGDQGNKASRWTRQEDVEEG
                     GPPGPREGPQSRPVAESTGQEATFPKATPLAQAAPLAEVDNPPTERDILPSDCAASAS
                     DSNTDHLDLGIEFSASAASGDELGLVEEKPAPCPSPEVLLPRLGWDDELQKPGAQVYM
                     HFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEEHKIETWREIYLQGCFKPLVSISPNDSLFEAVYALI
                     KNRIHRLPVLDPVSGAVLHILTHKRLLKFLHIFGTLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNETGQVVGLYSRFDVIHLAAQQTYNHLDMNV
                     GEALRQRTLCLEGVLSCQPHETLGEVIDRIVREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT      382 a    580 c    535 g    376 t
ORIGIN


  Query Match             8.6%;  Score 148;  DB 4;  Length 1873;
```

```
  Best Local Similarity   88.9%;  Pred. No. 3.2e-25;
  Matches  160;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy    1510 ccatcctaaccagggttccctgctgccccggccctccttcctctaccgcactatccaaga 1569
           ||  |  | |   ||  ||||||||||||||||||||||||||||||||||||||| ||||||||
Db     915 CCTGCACATCTTTGGCACCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACCATCCAAGA 974

Qy    1570 tttgggcatcggcacattccgagacttggctgtggtgctggagacagcacccatcctgac 1629
           |||||||||||||||||||||||||||||| |||||||||||  ||  || ||||||||||
Db     975 TTTGGGCATCGGCACATTCCGAGACTTGGCCGTGGTGCTGGAAACGGCGCCCATCCTGAC 1034

Qy    1630 tgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtgg 1689
           ||||||||||||||| |||||||||||||||||||| ||||||||||||||||||  |||
Db    1035 CGCACTGGACATCTTCGTGGACCGGCGTGTGTCTGCGCTGCCTGTGGTCAACGAAACTGG 1094


RESULT  14
AX099800
LOCUS       AX099800                1873 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 27 from Patent WO0120003.
ACCESSION   AX099800
VERSION     AX099800.1  GI:13538834
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 1873)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 27 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .1873
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
     CDS             1. .1395
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35800.1"
                     /db_xref="GI:13538835"
                     /translation="MSFLEQGESRSWPSRAVTTSSERSHGDQGNKASRWTRQEDVEEG
                     GPPGPREGPQSRPVAESTGQEATFPKATPLAQAAPLAEVDNPPTERDILPSDCAASAS
                     DSNTDHLDLGIEFSASAASGDELGLVEEKPAPCPSPEVLLPRLGWDDELQKPGAQVYM
                     HFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEEHKIETWREIYLQGCFKPLVSISPNDSLFEAVYALI
                     KNRIHRLPVLDPVSGAVLHILTHKRLLKFLHIFGTLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNETGQVVGLYSRFDVIHLAAQQTYNHLDMNV
                     GEALRQRTLCLEGVLSCQPHETLGEVIDRIVREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT      382 a    580 c    535 g    376 t
ORIGIN


  Query Match             8.6%;  Score 148;  DB 6;  Length 1873;
  Best Local Similarity   88.9%;  Pred. No. 3.2e-25;
  Matches  160;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy    1510 ccatcctaaccagggttccctgctgccccggccctccttcctctaccgcactatccaaga 1569
           ||  |  | |   ||  ||||||||||||||||||||||||||||||||||||||| ||||||||
Db     915 CCTGCACATCTTTGGCACCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACCATCCAAGA 974

Qy    1570 tttgggcatcggcacattccgagacttggctgtggtgctggagacagcacccatcctgac 1629
           |||||||||||||||||||||||||||||| |||||||||||  ||  || ||||||||||
Db     975 TTTGGGCATCGGCACATTCCGAGACTTGGCCGTGGTGCTGGAAACGGCGCCCATCCTGAC 1034

Qy    1630 tgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtgg 1689
           ||||||||||||||| |||||||||||||||||||| ||||||||||||||||||  |||
Db    1035 CGCACTGGACATCTTCGTGGACCGGCGTGTGTCTGCGCTGCCTGTGGTCAACGAAACTGG 1094


RESULT  15
AX099804
LOCUS       AX099804                2022 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 31 from Patent WO0120003.
ACCESSION   AX099804
VERSION     AX099804.1  GI:13538838
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 2022)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 31 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .2022
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
BASE COUNT      412 a    623 c    593 g    394 t
ORIGIN


  Query Match             8.6%;  Score 148;  DB 6;  Length 2022;
  Best Local Similarity   88.9%;  Pred. No. 3.1e-25;
  Matches  160;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy    1510 ccatcctaaccagggttccctgctgccccggccctccttcctctaccgcactatccaaga 1569
           ||  |  | |   ||  ||||||||||||||||||||||||||||||||||||||| ||||||||
Db    1065 CCTGCACATCTTTGGCACCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACCATCCAAGA 1124

Qy    1570 tttgggcatcggcacattccgagacttggctgtggtgctggagacagcacccatcctgac 1629
           |||||||||||||||||||||||||||||| |||||||||||  ||  || ||||||||||
Db    1125 TTTGGGCATCGGCACATTCCGAGACTTGGCCGTGGTGCTGGAAACGGCGCCCATCCTGAC 1184

Qy    1630 tgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtgg 1689
           ||||||||||||||| |||||||||||||||||||| ||||||||||||||||||  |||
Db    1185 CGCACTGGACATCTTCGTGGACCGGCGTGTGTCTGCGCTGCCTGTGGTCAACGAAACTGG 1244


Search completed: October  3, 2002, 14:46:08
Job time: 11888 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 12:32:10 ; Search time 719.93 Seconds
(without alignments)
4106.686 Million cell updates/sec

Title: US-09-826-581-3
Perfect score: 1722
Sequence: 1 cctggccccctcagatcaaga.....gatgagaggctcgggctgga 1722

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		1722	100.0	1722	22	AAH43683	PRKAG3 intron 4 -
c 2		351	20.4	378	22	ABA44706	Human breast cell
c 3		351	20.4	378	22	ABA55162	Human foetal liver
c 4		351	20.4	378	22	ABA24907	Probe #3373 for ge
c 5		351	20.4	378	22	AAK03423	Human brain expres
c 6		351	20.4	378	22	AAK28874	Human bone marrow
c 7		351	20.4	378	22	AAI13465	Probe #3398 for ge
c 8		351	20.4	378	22	AAI34821	Probe #3507 used t
c 9		351	20.4	378	22	AAI03344	Probe #3335 used t
10		196	11.4	547	22	ABA08485	Human AMP-activate
11		168.2	9.8	1647	22	AAH43685	PRKAG3 cDNA. Homo
12		168.2	9.8	2109	22	AAD03296	Human AMPK gamma s
13		168.2	9.8	2115	22	AAD03320	Human AMPK gamma s
14		148	8.6	1867	22	AAD03295	Pig AMPK gamma sub
15		148	8.6	1873	22	AAD03319	Pig AMPK gamma sub
16		148	8.6	2022	22	AAD03321	Sus scrofa PRKAG3
c 17		74.2	4.3	16525	22	AAK73303	Human immune/haema
18		70.4	4.1	602	22	AAH35203	Human colon cancer
19		70.4	4.1	1691	21	AAC98774	Human pancreatic c
20		68.8	4.0	1576	18	AAT85927	Mammalian AMPK-gam
21		68	3.9	39651	23	ABL18856	Drosophila melanog
22		65.8	3.8	3261	23	ABL18857	Drosophila melanog
c 23		65	3.8	92	22	ABA49850	Human breast cell
c 24		65	3.8	92	22	ABA67769	Human foetal liver
c 25		65	3.8	92	22	ABA34826	Probe #13292 for g
c 26		65	3.8	92	22	AAK16181	Human brain expres
c 27		65	3.8	92	22	AAK41922	Human bone marrow
c 28		65	3.8	92	22	AAI22692	Probe #12625 for g
c 29		65	3.8	92	22	AAI47988	Probe #16674 used
c 30		65	3.8	92	22	AAI08354	Probe #8345 used t
31		63.8	3.7	1435	20	AAX06882	Disease associated
32		63.8	3.7	1467	23	AAS84265	DNA encoding novel
33		63.8	3.7	2223	22	AAH14839	Human cDNA sequenc
34		62.6	3.6	350	21	AAC01661	Human secreted pro
35		57.2	3.3	735	22	AAH07561	Human cDNA clone (
36		49.8	2.9	2303	23	AAS84267	DNA encoding novel
37		49.4	2.9	151	24	ABI99496	Mouse ischaemic co
38		44	2.6	375	22	AAF66844	Novel human polynu
39		42.6	2.5	985	22	AAD08652	Mouse cancer assoc
40		42	2.4	61710	22	AAK83782	Human immune/haema
41		41.6	2.4	37664	22	AAK83781	Human immune/haema
c 42		40.8	2.4	1512	21	AAZ98024	Human secreted pro
c 43		40.8	2.4	1512	22	AAD11637	Human secreted pro
44		40.6	2.4	766	21	AAC44215	Arabidopsis thalia
c 45		40.6	2.4	8017	22	ABA19015	Human nervous syst

ALIGNMENTS

RESULT 1
AAH43683
ID AAH43683 standard; DNA; 1722 BP.
XX
AC AAH43683;
XX
DT 21-JAN-2002 (first entry)
XX
DE PRKAG3 intron 4 - intron 10.
XX
KW Human; AMP-activated protein kinase gamma 3 subunit; PRKAG3; variant;
KW metabolic disease; diabetes; obesity; substitution; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT intron 1..13
FT /*tag= a
FT /number = "Intron 4"
FT /note= "3' portion of intron 4"
FT exon 14..95
FT /*tag= b
FT /number= "Exon 5"
FT intron 96..552
FT /*tag= c
FT /number= "Intron 5"
FT exon 553..611
FT /*tag= d
FT /number= "Exon 6"
FT intron 612..736
FT /*tag= e
FT /number= "Intron 6"

```
FT   exon            737..782
FT                   /*tag= f
FT                   /number= "Exon 7"
FT   intron          783..986
FT                   /*tag= g
FT                   /number= "Intron 7"
FT   exon            987..1041
FT                   /*tag= h
FT                   /number= "Exon 8"
FT   intron          1042..1242
FT                   /*tag= i
FT                   /number= "Intron 8"
FT   exon            1243..1369
FT                   /*tag= j
FT                   /number= "Exon 9"
FT   intron          1370..1522
FT                   /*tag= k
FT                   /number= "Intron 9"
FT   exon            1523..1688
FT                   /*tag= l
FT                   /number= "Exon 10"
FT   intron          1689..1722
FT                   /*tag= i
FT                   /number= "Intron 10"
FT                   /note= "5' portion of intron 10"
XX
PN   WO200177305-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-SE00765.
XX
PR   07-APR-2000; 2000US-195665P.
XX
PA   (AREX-) AREXIS AB.
XX
PI   Andersson L, Luthman H, Marklund S;
XX
DR   WPI; 2001-657170/75.
XX
PT   New variants of human AMP-activated protein kinase gamma3 subunit
PT   associated with a metabolic disease e.g. diabetes or obesity and method
PT   for determining a risk estimate of diseases in subject by detecting the
PT   variant -
XX
PS   Example 1; Fig 3; 25pp; English.
XX
CC   The sequences given in AAH43681-84 represents genomic fragments
CC   encoding the human AMP-activated protein kinase gamma 3 subunit
CC   (PRKAG3). Detecting the presence of the PRKAG3 DNA, or a variant,
CC   is useful in determining a risk estimate of a metabolic disease,
CC   such as diabetes or obesity, in a subject. The variation may occur
CC   in exons 3, 4 or 10. In exon 3 variation may be a substitution of
CC   a G for a C at nucleotide 320, resulting in the amino acid
CC   substitution P71A; in exon 4 variation may be a substitution of a
CC   T for a C at nucleotide 550; and in exon 10 variation may be a
CC   substitution of a T for a C at nucleotide 1037, resulting in the
CC   amino acid substitution R340W. There may also be nucleotide variation
CC   in intron 6.
XX
SQ   Sequence 1722 BP; 321 A; 504 C; 534 G; 363 T; 0 other;
```

```
  Query Match             100.0%;  Score 1722;  DB 22;  Length 1722;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches 1722;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 cctggcccctcagatcaagaaggccttctttgctctggtggccaacggtgtgcgggcagc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 cctggcccctcagatcaagaaggccttctttgctctggtggccaacggtgtgcgggcagc 60

Qy      61 ccctctatgggacagcaagaagcagagctttgtgggtgaggagaggctggggaggtgaag 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 ccctctatgggacagcaagaagcagagctttgtgggtgaggagaggctggggaggtgaag 120

Qy     121 ggagatggaggaggtgaggggagatcttgtacggttgttctggggctgatctctgatat 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 ggagatggaggaggtgaggggagatcttgtacggttgttctggggctgatctctgatat 180

Qy     181 accacaagcttggcttcaggccaagcccagccaggggccagggtggaggaaagtccatcc 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 accacaagcttggcttcaggccaagcccagccaggggccagggtggaggaaagtccatcc 240

Qy     241 ggagtctgcatggccagctgggagaccctggggctcaatttccccatctgtggagccgct 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ggagtctgcatggccagctgggagaccctggggctcaatttccccatctgtggagccgct 300

Qy     301 atgaccagctgacacctttcacctccgctactgcatggccctgtgccataggtgctaggg 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 atgaccagctgacacctttcacctccgctactgcatggccctgtgccataggtgctaggg 360

Qy     361 agcaaatgggggggaggcaggagagaaagagccccacttctcaggcctgggggggctgcccc 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 agcaaatgggggggaggcaggagagaaagagccccacttctcaggcctgggggggctgcccc 420

Qy     421 actgtcctgttcccacagtccccactgtgtctcagcacaaggacactggcagggtgggga 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 actgtcctgttcccacagtccccactgtgtctcagcacaaggacactggcagggtgggga 480

Qy     481 ggggatctgaccctcaacctgccttccacccaaaggccccgggctgacctcctccccgcc 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 ggggatctgaccctcaacctgccttccacccaaaggccccgggctgacctcctccccgcc 540

Qy     541 cctcccctgcagggatgctgaccatcactgacttcatcctggtgctgcatcgctactaca 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 cctcccctgcagggatgctgaccatcactgacttcatcctggtgctgcatcgctactaca 600

Qy     601 ggtccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcgg 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 ggtccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcgg 660

Qy     661 aggggagtcctcctggagcctggtgccctagaagcccacgtctttctgacttctggagtc 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 aggggagtcctcctggagcctggtgccctagaagcccacgtctttctgacttctggagtc 720

Qy     721 ctgtcgatgtctctaggtccagatctatgagattgaacaacataagattgagacctggag 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 ctgtcgatgtctctaggtccagatctatgagattgaacaacataagattgagacctggag 780

Qy     781 gggtgagtggggagaggaacccggaaaggggctgttggtgatggtgggccagggcttaag 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 gggtgagtggggagaggaacccggaaaggggctgttggtgatggtgggccagggcttaag 840

Qy     841 gtggaggatgggcagtggggatgtcctggagtgaacaggggagggacaataggagcctcg 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 gtggaggatgggcagtggggatgtcctggagtgaacaggggagggacaataggagcctcg 900

Qy     901 ggtgcctgacggaagggaagctgcctgggactgcaaggtgaggcaggtgaccggctcccc 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 ggtgcctgacggaagggaagctgcctgggactgcaaggtgaggcaggtgaccggctcccc 960

Qy     961 tggcctgactctggctctttctgcagagatctacctgcaaggctgcttcaagcctctggt 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 tggcctgactctggctctttctgcagagatctacctgcaaggctgcttcaagcctctggt 1020

Qy    1021 ctccatctctcctaatgataggtgggtgtctctgctcattcacctgagcctcctcctccc 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 ctccatctctcctaatgataggtgggtgtctctgctcattcacctgagcctcctcctccc 1080

Qy    1081 acagtccccttccccagtcccactcagctctgaactcacctcttcatcctaggcggcaca 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 acagtccccttccccagtcccactcagctctgaactcacctcttcatcctaggcggcaca 1140

Qy    1141 cagacaagggagccttggtgccctgccctccttttttaggggcctgggatggaggttgtct 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 cagacaagggagccttggtgccctgccctccttttttaggggcctgggatggaggttgtct 1200
```

```
Qy   1201 ctccctaggctgccccgaggctcactgctcccatctctgcagcctgtttgaagctgtcta 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ctccctaggctgccccgaggctcactgctcccatctctgcagcctgtttgaagctgtcta 1260

Qy   1261 caccctcatcaagaaccggatccatcgcctgcctgttcttgacccggtgtcaggcaacgt 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 caccctcatcaagaaccggatccatcgcctgcctgttcttgacccggtgtcaggcaacgt 1320

Qy   1321 actccacatcctcacacacaaacgcctgctcaagttcctgcacatctttgtaagcctggg 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 actccacatcctcacacacaaacgcctgctcaagttcctgcacatctttgtaagcctggg 1380

Qy   1381 cccaggtgggaggaaggggagacctgggcaggtgatcagagggcctgaggagtcttcag 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 cccaggtgggaggaaggggagacctgggcaggtgatcagagggcctgaggagtcttcag 1440

Qy   1441 ccctagcagtcgtggggaagagctgggagccctcttgaagctgctggatccctgatctcc 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 ccctagcagtcgtggggaagagctgggagccctcttgaagctgctggatccctgatctcc 1500

Qy   1501 acctggtccccatcctaaccagggttccctgctgccccggccctccttcctctaccgcac 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 acctggtccccatcctaaccagggttccctgctgccccggccctccttcctctaccgcac 1560

Qy   1561 tatccaagatttgggcatcggcacattccgagacttggctgtggtgctggagacagcacc 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 tatccaagatttgggcatcggcacattccgagacttggctgtggtgctggagacagcacc 1620

Qy   1621 catcctgactgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaa 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 catcctgactgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaa 1680

Qy   1681 cgaatgtggtacccacccccaggatgagaggctcgggctgga 1722
          |||||||||||||||||||||||||||||||||||||||||
Db   1681 cgaatgtggtacccacccccaggatgagaggctcgggctgga 1722
```

```
RESULT   2
ABA44706/c
ID   ABA44706 standard; DNA; 378 BP.
XX
AC   ABA44706;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human breast cell single exon nucleic acid probe #3401.
XX
KW   Human; microarray; single exon probe; gene expression; breast;
KW   disease; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157271-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00662.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-496933/54.
XX
PT   New spatially-addressable set of single exon nucleic acid probes,
PT   useful for measuring gene expression in sample derived from human
PT   breast, comprises number of single exon nucleic acid probes  -
XX
PS   Claim 1; SEQ ID NO 3401; 327pp + sequence listing; English.
XX
CC   The invention relates to a spatially-addressable set of single exon
CC   nucleic acid probes for measuring gene expression in a sample derived
CC   from human breast and BT 474 cells. The method involves contacting
CC   the probes with a collection of detectably labelled nucleic acids
CC   derived from mRNA of human breast, and then measuring the label
CC   bound to each probe of the microarray. The probes are useful for
CC   verifying the expression of regions of genomic DNA predicted to
CC   encode proteins. They are useful for gene discovery, and for
CC   determining predisposition and/or prognosing breast disease. Gene
CC   expression analysis is useful for assessing the toxicity of chemical
CC   agents on cells. The microarray of this invention presents a far greater
CC   diversity of probes for measuring gene expression, with far less bias
CC   than expressed sequence tag microarrays. The method is suitable for
CC   rapid production of functional information from genomic sequence. The
CC   present sequence is a single exon nucleic acid probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 378 BP; 80 A; 128 C; 97 G; 73 T; 0 other;
```

```
  Query Match             20.4%;  Score 351;  DB 22;  Length 378;
  Best Local Similarity   99.7%;  Pred. No. 6.9e-83;
  Matches  362;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1 cctggcccctcagatcaagaaggccttctttgctctggtggccaacggtgtgcgggcagc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     362 CCTGGCCCCTCAGATCAAGAAGGCCTTCTTTGCTCTGGTGGCCAACGGTGTGCGGGCAGC 303

Qy      61 ccctctatgggacagcaagaagcagagctttgtgggtgaggagaggctggggaggtgaag 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302 CCCTCTATGGGACAGCAAGAAGCAGAGCTTTGTGGGTGAGGAGAGGCTGGGGAGGTGAAG 243

Qy     121 ggagatggaggaggtgaggggagatcttgtacggttgttctggggctgatctctgatat 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     242 GGAGATGGAGGAGGTGAGGGGGAGATCTTGTACGGTTGTTCTGGGGCTGATCTCTGATAT 183

Qy     181 accacaagcttggcttcaggccaagcccagccaggggccagggtggaggaaagtccatcc 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 ACCACAAGCTTGGCTTCAGGCCAAGCCCAGCCAGGGGCCAGGGTGGAGGAAAGTCCATCC 123

Qy     241 ggagtctgcatggccagctgggagaccctggggctcaatttccccatctgtggagccgct 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 GGAGTCTGCATGGCCAGCTGGGAGACCCTGGGGCTCAATTTCCCCATCTGTGGAGCCGCT 63

Qy     301 atgaccagctgacacctttcacctccgctactgcatggccctgtgccataggtgctaggg 360
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db      62 ATGACCAGCTGACACCTTTCACCTCCGCTACTGCATGGCCCTGTG-CATAGGTGCTAGGG 4

Qy     361 agc 363
           |||
Db       3 AGC 1
```

```
RESULT   3
ABA55162/c
ID   ABA55162 standard; DNA; 378 BP.
XX
AC   ABA55162;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #3467.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
```

```
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00669.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human fetal liver -
XX
PS   Claim 1; SEQ ID NO 3467; 639pp + sequence listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for
CC   measuring human gene expression in a sample derived from human foetal
CC   liver. The single exon nucleic acid probes may be used for predicting,
CC   measuring and displaying gene expression in samples derived from human
CC   fetal liver. The present sequence is a single exon nucleic acid
CC   probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 378 BP; 80 A; 128 C; 97 G; 73 T; 0 other;


  Query Match             20.4%;  Score 351;  DB 22;  Length 378;
  Best Local Similarity   99.7%;  Pred. No. 6.9e-83;
  Matches  362;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1 cctggccccctcagatcaagaaggccttctttgctctggtggccaacggtgtgcgggcagc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     362 CCTGGCCCCTCAGATCAAGAAGGCCTTCTTTGCTCTGGTGGCCAACGGTGTGCGGGCAGC 303

Qy      61 ccctctatgggacagcaagaagcagagctttgtgggtgaggagaggctggggaggtgaag 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302 CCCTCTATGGGACAGCAAGAAGCAGAGCTTTGTGGGTGAGGAGAGGCTGGGGAGGTGAAG 243

Qy     121 ggagatggaggaggtgagggggagatcttgtacggttgttctggggctgatctctgatat 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     242 GGAGATGGAGGAGGTGAGGGGGAGATCTTGTACGGTTGTTCTGGGGCTGATCTCTGATAT 183

Qy     181 accacaagcttggcttcaggccaagcccagccaggggccagggtggaggaaagtccatcc 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 ACCACAAGCTTGGCTTCAGGCCAAGCCCAGCCAGGGGCCAGGGTGGAGGAAAGTCCATCC 123

Qy     241 ggagtctgcatggccagctgggagaccctggggctcaatttccccatctgtggagccgct 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 GGAGTCTGCATGGCCAGCTGGGAGACCCTGGGGCTCAATTTCCCCATCTGTGGAGCCGCT 63

Qy     301 atgaccagctgacacctttcacctccgctactgcatggccctgtgccataggtgctaggg 360
           |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db      62 ATGACCAGCTGACACCTTTCACCTCCGCTACTGCATGGCCCTGTG-CATAGGTGCTAGGG 4

Qy     361 agc 363
           |||
Db       3 AGC 1
```

```
RESULT   4
ABA24907/c
ID   ABA24907 standard; DNA; 378 BP.
XX
AC   ABA24907;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #3373 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00666.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts -
XX
PS   Claim 1; SEQ ID No 3373; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 378 BP; 80 A; 128 C; 97 G; 73 T; 0 other;


  Query Match             20.4%;  Score 351;  DB 22;  Length 378;
  Best Local Similarity   99.7%;  Pred. No. 6.9e-83;
  Matches  362;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1 cctggcccctcagatcaagaaggccttctttgctctggtggccaacggtgtgcgggcagc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     362 CCTGGCCCCTCAGATCAAGAAGGCCTTCTTTGCTCTGGTGGCCAACGGTGTGCGGGCAGC 303

Qy      61 ccctctatgggacagcaagaagcagagctttgtgggtgaggagaggctggggaggtgaag 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302 CCCTCTATGGGACAGCAAGAAGCAGAGCTTTGTGGGTGAGGAGAGGCTGGGGAGGTGAAG 243

Qy     121 ggagatggaggaggtgagggggagatcttgtacggttgttctggggctgatctctgatat 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     242 GGAGATGGAGGAGGTGAGGGGGAGATCTTGTACGGTTGTTCTGGGGCTGATCTCTGATAT 183
```

```
Qy    181 accacaagcttggcttcaggccaagcccagccagggccagggtggaggaaagtccatcc 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    182 ACCACAAGCTTGGCTTCAGGCCAAGCCCAGCCAGGGGCCAGGGTGGAGGAAAGTCCATCC 123

Qy    241 ggagtctgcatggccagctgggagaccctggggctcaatttccccatctgtggagccgct 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122 GGAGTCTGCATGGCCAGCTGGGAGACCCTGGGGCTCAATTTCCCCATCTGTGGAGCCGCT 63

Qy    301 atgaccagctgacacctttcacctccgctactgcatggccctgtgccataggtgctaggg 360
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     62 ATGACCAGCTGACACCTTTCACCTCCGCTACTGCATGGCCCTGTG-CATAGGTGCTAGGG 4

Qy    361 agc 363
          |||
Db      3 AGC 1


RESULT   5
AAK03423/c
ID   AAK03423 standard; DNA; 378 BP.
XX
AC   AAK03423;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 3414.
XX
KW   Human; brain expressed exon; gene expression analysis; probe;
KW   microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW   epilepsy; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00667.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains -
XX
PS   Example 4; SEQ ID NO: 3414; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention.
XX
SQ   Sequence 378 BP; 80 A; 128 C; 97 G; 73 T; 0 other;


  Query Match             20.4%;  Score 351;  DB 22;  Length 378;
  Best Local Similarity   99.7%;  Pred. No. 6.9e-83;
  Matches  362;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy      1 cctggcccctcagatcaagaaggccttctttgctctggtggccaacggtgtgcgggcagc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    362 CCTGGCCCCTCAGATCAAGAAGGCCTTCTTTGCTCTGGTGGCCAACGGTGTGCGGGCAGC 303

Qy     61 ccctctatgggacagcaagaagcagagctttgtgggtgaggagaggctggggaggtgaag 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    302 CCCTCTATGGGACAGCAAGAAGCAGAGCTTTGTGGGTGAGGAGAGGCTGGGGAGGTGAAG 243

Qy    121 ggagatggaggaggtgagggggagatcttgtacggttgttctggggctgatctctgatat 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    242 GGAGATGGAGGAGGTGAGGGGGAGATCTTGTACGGTTGTTCTGGGGCTGATCTCTGATAT 183

Qy    181 accacaagcttggcttcaggccaagcccagccaggggccagggtggaggaaagtccatcc 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    182 ACCACAAGCTTGGCTTCAGGCCAAGCCCAGCCAGGGGCCAGGGTGGAGGAAAGTCCATCC 123

Qy    241 ggagtctgcatggccagctgggagaccctggggctcaatttccccatctgtggagccgct 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    122 GGAGTCTGCATGGCCAGCTGGGAGACCCTGGGGCTCAATTTCCCCATCTGTGGAGCCGCT 63

Qy    301 atgaccagctgacacctttcacctccgctactgcatggccctgtgccataggtgctaggg 360
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db     62 ATGACCAGCTGACACCTTTCACCTCCGCTACTGCATGGCCCTGTG-CATAGGTGCTAGGG 4

Qy    361 agc 363
          |||
Db      3 AGC 1


RESULT   6
AAK28874/c
ID   AAK28874 standard; DNA; 378 BP.
XX
AC   AAK28874;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 3431.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00668.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human bone marrow -
XX
PS   Example 4; SEQ ID NO: 3431; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
```

```
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention.
XX
SQ   Sequence 378 BP; 80 A; 128 C; 97 G; 73 T; 0 other;


  Query Match             20.4%;  Score 351;  DB 22;  Length 378;
  Best Local Similarity   99.7%;  Pred. No. 6.9e-83;
  Matches  362;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1 cctggcccctcagatcaagaaggccttctttgctctggtggccaacggtgtgcgggcagc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     362 CCTGGCCCCTCAGATCAAGAAGGCCTTCTTTGCTCTGGTGGCCAACGGTGTGCGGGCAGC 303

Qy      61 ccctctatgggacagcaagaagcagagctttgtgggtgaggagaggctggggaggtgaag 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302 CCCTCTATGGGACAGCAAGAAGCAGAGCTTTGTGGGTGAGGAGAGGCTGGGGAGGTGAAG 243

Qy     121 ggagatggaggaggtgaggggggagatcttgtacggttgttctggggctgatctctgatat 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     242 GGAGATGGAGGAGGTGAGGGGGAGATCTTGTACGGTTGTTCTGGGGCTGATCTCTGATAT 183

Qy     181 accacaagcttggcttcaggccaagcccagccagggggccagggtggaggaaagtccatcc 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 ACCACAAGCTTGGCTTCAGGCCAAGCCCAGCCAGGGGCCAGGGTGGAGGAAAGTCCATCC 123

Qy     241 ggagtctgcatggccagctgggagaccctggggctcaatttccccatctgtggagccgct 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 GGAGTCTGCATGGCCAGCTGGGAGACCCTGGGGCTCAATTTCCCCATCTGTGGAGCCGCT 63

Qy     301 atgaccagctgacacctttcacctccgctactgcatggccctgtgccataggtgctaggg 360
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db      62 ATGACCAGCTGACACCTTTCACCTCCGCTACTGCATGGCCCTGTG-CATAGGTGCTAGGG 4

Qy     361 agc 363
           |||
Db       3 AGC 1


RESULT   7
AAI13465/c
ID   AAI13465 standard; DNA; 378 BP.
XX
AC   AAI13465;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #3398 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID No 3398; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 378 BP; 80 A; 128 C; 97 G; 73 T; 0 other;


  Query Match             20.4%;  Score 351;  DB 22;  Length 378;
  Best Local Similarity   99.7%;  Pred. No. 6.9e-83;
  Matches  362;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1 cctggcccctcagatcaagaaggccttctttgctctggtggccaacggtgtgcgggcagc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     362 CCTGGCCCCTCAGATCAAGAAGGCCTTCTTTGCTCTGGTGGCCAACGGTGTGCGGGCAGC 303

Qy      61 ccctctatgggacagcaagaagcagagctttgtgggtgaggagaggctggggaggtgaag 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302 CCCTCTATGGGACAGCAAGAAGCAGAGCTTTGTGGGTGAGGAGAGGCTGGGGAGGTGAAG 243

Qy     121 ggagatggaggaggtgaggggggagatcttgtacggttgttctggggctgatctctgatat 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     242 GGAGATGGAGGAGGTGAGGGGGAGATCTTGTACGGTTGTTCTGGGGCTGATCTCTGATAT 183

Qy     181 accacaagcttggcttcaggccaagcccagccaggggccagggtggaggaaagtccatcc 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 ACCACAAGCTTGGCTTCAGGCCAAGCCCAGCCAGGGGCCAGGGTGGAGGAAAGTCCATCC 123

Qy     241 ggagtctgcatggccagctgggagaccctggggctcaatttccccatctgtggagccgct 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 GGAGTCTGCATGGCCAGCTGGGAGACCCTGGGGCTCAATTTCCCCATCTGTGGAGCCGCT 63

Qy     301 atgaccagctgacacctttcacctccgctactgcatggccctgtgccataggtgctaggg 360
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db      62 ATGACCAGCTGACACCTTTCACCTCCGCTACTGCATGGCCCTGTG-CATAGGTGCTAGGG 4

Qy     361 agc 363
           |||
Db       3 AGC 1


RESULT   8
AAI34821/c
ID   AAI34821 standard; DNA; 378 BP.
XX
AC   AAI34821;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #3507 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00663.
```

```
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human placenta -
XX
PS   Claim 25; SEQ ID No 3507; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders.
XX
SQ   Sequence 378 BP; 80 A; 128 C; 97 G; 73 T; 0 other;


  Query Match             20.4%;  Score 351;  DB 22;  Length 378;
  Best Local Similarity   99.7%;  Pred. No. 6.9e-83;
  Matches  362;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1 cctggcccctcagatcaagaaggccttctttgctctggtggccaacggtgtgcgggcagc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     362 CCTGGCCCCTCAGATCAAGAAGGCCTTCTTTGCTCTGGTGGCCAACGGTGTGCGGGCAGC 303

Qy      61 ccctctatgggacagcaagaagcagagctttgtgggtgaggagaggctggggaggtgaag 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302 CCCTCTATGGGACAGCAAGAAGCAGAGCTTTGTGGGTGAGGAGAGGCTGGGGAGGTGAAG 243

Qy     121 ggagatggaggaggtgagggggagatcttgtacggttgttctggggctgatctctgatat 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     242 GGAGATGGAGGAGGTGAGGGGGAGATCTTGTACGGTTGTTCTGGGGCTGATCTCTGATAT 183

Qy     181 accacaagcttggcttcaggccaagcccagccaggggccagggtggaggaaagtccatcc 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 ACCACAAGCTTGGCTTCAGGCCAAGCCCAGCCAGGGGCCAGGGTGGAGGAAAGTCCATCC 123

Qy     241 ggagtctgcatggccagctgggagaccctggggctcaatttccccatctgtggagccgct 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 GGAGTCTGCATGGCCAGCTGGGAGACCCTGGGGCTCAATTTCCCCATCTGTGGAGCCGCT 63

Qy     301 atgaccagctgacacctttcacctccgctactgcatggccctgtgccataggtgctaggg 360
           |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db      62 ATGACCAGCTGACACCTTTCACCTCCGCTACTGCATGGCCCTGTG-CATAGGTGCTAGGG 4

Qy     361 agc 363
           |||
Db       3 AGC 1


RESULT   9
AAI03344/c
ID   AAI03344 standard; DNA; 378 BP.
XX
AC   AAI03344;
XX
DT   09-OCT-2001  (first entry)
XX
DE   Probe #3335 used to measure gene expression in human breast sample.
XX
KW   Probe; human; breast disease; breast cancer; development disorder; ss;
KW   inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS   Homo sapiens.
XX
PN   WO200157270-A2.
XX
PD   09-AUG-2001.
XX
PF   29-JAN-2001; 2001WO-US00661.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-476286/51.
XX
PT   Novel single exon nucleic acid probe used to measuring gene expression
PT   in a human breast -
XX
PS   Claim 25; SEQ ID No 3335; 322pp; English.
XX
CC   The present invention relates to novel single exon nucleic acid probes.
CC   The present sequence is one such probe. The probes are useful for
CC   measuring human gene expression in a human breast sample, where the probe
CC   hybridises at high stringency to a nucleic acid expressed in the human
CC   breast. The probes are useful for predicting, diagnosing, grading,
CC   staging, monitoring and prognosing diseases of the human breast,
CC   particularly those diseases with polygenic aetiology. The diseases
CC   include: breast cancer, disorders of development, inflammatory diseases
CC   of the breast, fibrocystic changes, proliferative breast disease and
CC   non-carcinoma tumours.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 378 BP; 80 A; 128 C; 97 G; 73 T; 0 other;


  Query Match             20.4%;  Score 351;  DB 22;  Length 378;
  Best Local Similarity   99.7%;  Pred. No. 6.9e-83;
  Matches  362;  Conservative    0;  Mismatches    0;  Indels    1;  Gaps    1;

Qy       1 cctggcccctcagatcaagaaggccttctttgctctggtggccaacggtgtgcgggcagc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     362 CCTGGCCCCTCAGATCAAGAAGGCCTTCTTTGCTCTGGTGGCCAACGGTGTGCGGGCAGC 303

Qy      61 ccctctatgggacagcaagaagcagagctttgtgggtgaggagaggctggggaggtgaag 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     302 CCCTCTATGGGACAGCAAGAAGCAGAGCTTTGTGGGTGAGGAGAGGCTGGGGAGGTGAAG 243

Qy     121 ggagatggaggaggtgagggggagatcttgtacggttgttctggggctgatctctgatat 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     242 GGAGATGGAGGAGGTGAGGGGGAGATCTTGTACGGTTGTTCTGGGGCTGATCTCTGATAT 183

Qy     181 accacaagcttggcttcaggccaagcccagccaggggccagggtggaggaaagtccatcc 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     182 ACCACAAGCTTGGCTTCAGGCCAAGCCCAGCCAGGGGCCAGGGTGGAGGAAAGTCCATCC 123

Qy     241 ggagtctgcatggccagctgggagaccctggggctcaatttccccatctgtggagccgct 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     122 GGAGTCTGCATGGCCAGCTGGGAGACCCTGGGGCTCAATTTCCCCATCTGTGGAGCCGCT 63

Qy     301 atgaccagctgacacctttcacctccgctactgcatggccctgtgccataggtgctaggg 360
           |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db      62 ATGACCAGCTGACACCTTTCACCTCCGCTACTGCATGGCCCTGTG-CATAGGTGCTAGGG 4
```

```
Qy     361 agc 363
           |||
Db       3 AGC 1


RESULT  10
ABA08485
ID   ABA08485 standard; cDNA; 547 BP.
XX
AC   ABA08485;
XX
DT   11-JAN-2002  (first entry)
XX
DE   Human AMP-activated protein kinase subunit homologue cDNA, SEQ ID NO:261.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW   haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW   inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW   proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW   myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW   chronic inflammatory condition; proliferative retinopathy;
KW   atherosclerosis; coronary heart disease; arterial ischaemia;
KW   bone disorder; osteoporosis; vascular growth disorder;
KW   tissue regeneration; wound healing; infection; immune disorder;
KW   cell culture; drug screening; gene therapy; antiinflammatory;
KW   antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW   cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW   antifungal; vulnerary; antiulcer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157188-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US03800.
XX
PR   03-FEB-2000; 2000US-0496914.
PR   27-APR-2000; 2000US-0560875.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-457740/49.
DR   P-PSDB; ABB11241.
XX
PT   Human proteins and DNA encoding sequences useful for preventing,
PT   treating or ameliorating a medical condition in a mammalian subject
PT   e.g. arthritis and cancer -
XX
PS   Claim 1; Page 429; 1963pp; English.
XX
CC   Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC   sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC   invention also relates to vectors and recombinant host cells comprising a
CC   nucleotide of the invention, methods of producing the novel polypeptides,
CC   antibodies against the polypeptides, methods of detecting the nucleotides
CC   or polypeptides in a sample, and methods of identifying compounds which
CC   bind to polypeptides of the invention. Although novel, many of the
CC   polypeptides of the invention have homology to known proteins, thereby
CC   giving an insight into their probable biological activities, and hence
CC   potential therapeutic applications. The polypeptides of the invention may
CC   have various activities, including cytokine, cell proliferation or cell
CC   differentiation activities; stem cell growth factor activity;
CC   haematopoiesis regulatory activity; tissue growth activity;
CC   immunomodulatory activity; activin- or inhibin-related activities;
CC   chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC   thrombolytic activities; receptor or ligand activities; or may be
CC   involved in oncogenesis, cancer cell proliferation or metastasis.
CC   Depending on their biological activities, polypeptides and nucleotides of
CC   the invention are useful for preventing, treating or ameliorating medical
CC   conditions, e.g., by protein or gene therapy. Such conditions include
CC   cancers, haematopoietic disorders (e.g.,myeloid or lymphoid cell
CC   disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC   proliferative retinopathy, atherosclerosis, coronary heart disease,
CC   arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC   vascular growth. Polypeptides involved with tissue regeneration and
CC   repair (or nucleic acids encoding them) may be used to promote wound
CC   healing (e.g., of burns, incisions and ulcers), while those with
CC   immunomodulatory activities may be used in the treatment of viral,
CC   bacterial and fungal infections in addition to immune disorders.
CC   Polypeptides with growth factor activity may be used in cell cultures to
CC   promote cell growth. For example, such polypeptides may be used to
CC   manipulate stem cells in culture to give rise to neuroepithelial cells
CC   that can be used to augment or replace cells damaged by illness,
CC   autoimmune disease or accidental damage. The polypeptides and nucleotides
CC   may also be used in the diagnosis of the above conditions, and in drug
CC   screening techniques. The present sequence represents a cDNA encoding a
CC   novel human polypeptide of the invention.
XX
SQ   Sequence 547 BP; 112 A; 172 C; 133 G; 130 T; 0 other;


  Query Match             11.4%;  Score 196;  DB 22;  Length 547;
  Best Local Similarity   95.3%;  Pred. No. 7.3e-42;
  Matches  202;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy    1510 ccatcctaaccagggttccctgctgccccggccctccttcctctaccgcactatccaaga 1569
           ||  |  | |   ||||||||||||||||||||||||||||||||||||||||||||||
Db     272 cctgcacatctttggttccctgctgccccggccctccttcctctaccgcactatccaaga 331

Qy    1570 tttgggcatcggcacattccgagacttggctgtggtgctggagacagcacccatcctgac 1629
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     332 tttgggcatcggcacattccgagacttggctgtggtgctggagacagcacccatcctgac 391

Qy    1630 tgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtgg 1689
           |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
Db     392 tgcactggacatctttgtggaccggcgtgtgtctgcactggctgtggtcaacgaatgtgg 451

Qy    1690 tacccacccccaggatgagaggctcgggctgg 1721
           |||||||||||| ||||||||||||||||||
Db     452 tacccacccccaagatgagaggctcgggctgg 483


RESULT  11
AAH43685
ID   AAH43685 standard; cDNA; 1647 BP.
XX
AC   AAH43685;
XX
DT   21-JAN-2002  (first entry)
XX
DE   PRKAG3 cDNA.
XX
KW   Human; AMP-activated protein kinase gamma 3 subunit; PRKAG3; variant;
KW   metabolic disease; diabetes; obesity; substitution; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             20..1489
FT                   /*tag=  a
FT                   /product= "PRKAG3"
FT   variation       230
FT                   /*tag=  b
FT                   /label= "C230G"
FT                   /note= "Causes P71A"
FT   variation       559
FT                   /*tag=  c
FT                   /label= "T559C"
FT                   /note= "Silent variation"
FT   variation       1037
FT                   /*tag=  d
FT                   /label= "C1037T"
```

```
FT                   /note= "Causes R340W"
XX
PN   WO200177305-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-SE00765.
XX
PR   07-APR-2000; 2000US-195665P.
XX
PA   (AREX-) AREXIS AB.
XX
PI   Andersson L,  Luthman H,  Marklund S;
XX
DR   WPI; 2001-657170/75.
DR   P-PSDB; QQB47679.
XX
PT   New variants of human AMP-activated protein kinase gamma3 subunit
PT   associated with a metabolic disease e.g. diabetes or obesity and method
PT   for determining a risk estimate of diseases in subject by detecting the
PT   variant -
XX
PS   Disclosure; Fig 5; 25pp; English.
XX
CC   This sequence represents the full length cDNA encoding the human
CC   AMP-activated protein kinase gamma 3 subunit (PRKAG3). Detecting
CC   the presence of the PRKAG3 DNA, or a variant, is useful in determining
CC   a risk estimate of a metabolic disease, such as diabetes or obesity,
CC   in a subject.  The variation may occur in exons 3, 4 or 10.  In exon
CC   3 variation may be a substitution of a G for a C at nucleotide 320,
CC   resulting in the amino acid substitution P71A; in exon 4 variation may
CC   be a substitution of a T for a C at nucleotide 550; and in exon 10
CC   variation may be a substitution of a T for a C at nucleotide 1037,
CC   resulting in the amino acid substitution R340W.  There may also be
CC   nucleotide variation in intron 6.  The numbering of these
CC   variations is based on the full length cDNA as given, rather than on
CC   position 1 of the open reading frame.
XX
SQ   Sequence 1647 BP; 346 A; 502 C; 462 G; 337 T; 0 other;
```

```
  Query Match             9.8%;  Score 168.2;  DB 22;  Length 1647;
  Best Local Similarity   95.6%;  Pred. No. 2.5e-34;
  Matches  173;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    1510 ccatcctaaccagggttccctgctgccccggccctccttcctctaccgcactatccaaga 1569
           ||  |  | |   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1009 cctgcacatctttggttccctgctgccccggccctccttcctctaccgcactatccaaga 1068

Qy    1570 tttgggcatcggcacattccgagacttggctgtggtgctggagacagcacccatcctgac 1629
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1069 tttgggcatcggcacattccgagacttggctgtggtgctggagacagcacccatcctgac 1128

Qy    1630 tgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtgg 1689
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1129 tgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtgg 1188

Qy    1690 t 1690
           |
Db    1189 t 1189
```

```
RESULT  12
AAD03296
ID   AAD03296 standard; DNA; 2109 BP.
XX
AC   AAD03296;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Human AMPK gamma subunit muscle-specific isoform, PRKAG3 cDNA.
XX
KW   Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW   PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW   genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW   cystathione beta synthase; CBS; cardiant; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1..471
FT                   /*tag= a
FT   CDS             472..1389
FT                   /*tag= b
FT                   /product= "Human Prkag3 protein"
FT   3'UTR           1390..2109
FT                   /*tag= c
XX
PN   WO200120003-A2.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-EP09896.
XX
PR   10-SEP-1999;   99EP-0402236.
PR   18-MAY-2000; 2000EP-0401388.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA   (ANDE/) ANDERSSON L.
PA   (LOOF/) LOOFT C.
PA   (KALM/) KALM E.
XX
PI   Andersson L,  Looft C,  Kalm E,  Milan D,  Robic A,  Rogel-Gaillard C;
PI   Iannuccelli N,  Gellin J,  Le Roy P,  Chardon P;
XX
DR   WPI; 2001-244810/25.
DR   P-PSDB; AAE00221.
XX
PT   New variants of the gamma subunit of vertebrate adenosine
PT   monophosphate-activated kinase for diagnosis or treatment of disorders
PT   associated with energy metabolism such as diabetes, obesity, and
PT   myopathy  -
XX
PS   Claim 12; Fig 2; 71pp; English.
XX
CC   The present sequence is a cDNA encoding human adenosine monophosphate
CC   (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC   PRKAG3. Mutation in Prkag3 results in an altered regulation of
CC   carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is
CC   useful as therapeutic for treating carbohydrate metabolism disorders such
CC   as diabetes, obesity, and disorders associated with muscle metabolism
CC   such as myopathy and cardiovascular diseases, to modulate AMPK
CC   activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC   and its functionally altered mutants are useful for the diagnostic
CC   evaluation, genetic testing and prognosis of a metabolic disorder,
CC   preferably a carbohydrate metabolism disorder. Primers that can detect
CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
CC   from the expression of a functionally altered allele of PRKAG3.
CC   Transgenic animal and host cell transformed with PRKAG3 or a
CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC   screening compounds able to modulate AMPK activity. Nucleic acid
CC   encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
CC   of PRKAG3 and is useful in gene therapy.
XX
SQ   Sequence 2109 BP; 458 A; 621 C; 560 G; 470 T; 0 other;
```

```
  Query Match             9.8%;  Score 168.2;  DB 22;  Length 2109;
  Best Local Similarity   95.6%;  Pred. No. 2.8e-34;
  Matches  173;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    1510 ccatcctaaccagggttccctgctgccccggccctccttcctctaccgcactatccaaga 1569
           ||  |  | |   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     909 cctgcacatctttggttccctgctgccccggccctccttcctctaccgcactatccaaga 968
```

```
Qy   1570 tttgggcatcggcacattccgagacttggctgtggtgctggagacagcacccatcctgac 1629
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    969 tttgggcatcggcacattccgagacttggctgtggtgctggagacagcacccatcctgac 1028

Qy   1630 tgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtgg 1689
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1029 tgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtgg 1088

Qy   1690 t 1690
          |
Db   1089 t 1089
```

```
RESULT  13
AAD03320
ID   AAD03320 standard; cDNA; 2115 BP.
XX
AC   AAD03320;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Human AMPK gamma subunit muscle-specific isoform, complete PRKAG3 cDNA.
XX
KW   Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW   PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW   genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW   cystathione beta synthase; CBS; cardiant; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1395
FT                   /*tag= a
FT                   /product= "Human complete Prkag3 protein"
XX
PN   WO200120003-A2.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-EP09896.
XX
PR   10-SEP-1999;   99EP-0402236.
PR   18-MAY-2000; 2000EP-0401388.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA   (ANDE/) ANDERSSON L.
PA   (LOOF/) LOOFT C.
PA   (KALM/) KALM E.
XX
PI   Andersson L,  Looft C,  Kalm E,  Milan D,  Robic A,  Rogel-Gaillard C;
PI   Iannuccelli N,  Gellin J,  Le Roy P,  Chardon P;
XX
DR   WPI; 2001-244810/25.
DR   P-PSDB; AAE00223.
XX
PT   New variants of the gamma subunit of vertebrate adenosine
PT   monophosphate-activated kinase for diagnosis or treatment of disorders
PT   associated with energy metabolism such as diabetes, obesity, and
PT   myopathy  -
XX
PS   Claim 12; Page 65-68; 71pp; English.
XX
CC   The present sequence is a cDNA encoding human adenosine monophosphate
CC   (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC   complete PRKAG3. Mutation in Prkag3 results in an altered regulation of
CC   carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is
CC   useful as therapeutic for treating carbohydrate metabolism disorders such
CC   as diabetes, obesity, and disorders associated with muscle metabolism
CC   such as myopathy and cardiovascular diseases, to modulate AMPK
CC   activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC   and its functionally altered mutants are useful for the diagnostic
CC   evaluation, genetic testing and prognosis of a metabolic disorder,
CC   preferably a carbohydrate metabolism disorder. Primers that can detect
CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
CC   from the expression of a functionally altered allele of PRKAG3.
CC   Transgenic animal and host cell transformed with PRKAG3 or a
CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC   screening compounds able to modulate AMPK activity. Nucleic acid
CC   encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
CC   of PRKAG3 and is useful in gene therapy.
XX
SQ   Sequence 2115 BP; 460 A; 622 C; 562 G; 471 T; 0 other;
```

```
  Query Match             9.8%;  Score 168.2;  DB 22;  Length 2115;
  Best Local Similarity   95.6%;  Pred. No. 2.8e-34;
  Matches  173;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   1510 ccatcctaaccagggttccctgctgccccggccctccttcctctaccgcactatccaaga 1569
          ||  |  | |   ||||||||||||||||||||||||||||||||||||||||||||||
Db    915 cctgcacatctttggttccctgctgccccggccctccttcctctaccgcactatccaaga 974

Qy   1570 tttgggcatcggcacattccgagacttggctgtggtgctggagacagcacccatcctgac 1629
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    975 tttgggcatcggcacattccgagacttggctgtggtgctggagacagcacccatcctgac 1034

Qy   1630 tgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtgg 1689
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1035 tgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtgg 1094

Qy   1690 t 1690
          |
Db   1095 t 1095
```

```
RESULT  14
AAD03295
ID   AAD03295 standard; cDNA; 1867 BP.
XX
AC   AAD03295;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Pig AMPK gamma subunit muscle-specific isoform, PRKAG3 cDNA.
XX
KW   Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW   PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW   genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW   cystathione beta synthase; CBS; cardiant; gene therapy; RN locus;
KW   chromosome 15; ss.
XX
OS   Sus scrofa.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1..471
FT                   /*tag= a
FT   CDS             472..1389
FT                   /*tag= b
FT                   /product= "Sus scrofa PRKAG3 protein"
FT   3'UTR           1390..1867
FT                   /*tag= c
XX
PN   WO200120003-A2.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-EP09896.
XX
PR   10-SEP-1999;   99EP-0402236.
PR   18-MAY-2000; 2000EP-0401388.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA   (ANDE/) ANDERSSON L.
```

```
PA   (LOOF/) LOOFT C.
PA   (KALM/) KALM E.
XX
PI   Andersson L,  Looft C,  Kalm E,  Milan D,  Robic A,  Rogel-Gaillard C;
PI   Iannuccelli N,  Gellin J,  Le Roy P,  Chardon P;
XX
DR   WPI; 2001-244810/25.
DR   P-PSDB; AAE00220.
XX
PT   New variants of the gamma subunit of vertebrate adenosine
PT   monophosphate-activated kinase for diagnosis or treatment of disorders
PT   associated with energy metabolism such as diabetes, obesity, and
PT   myopathy  -
XX
PS   Claim 12; Fig 2; 71pp; English.
XX
CC   The present sequence is a cDNA encoding pig adenosine monophosphate
CC   (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC   PRKAG3. Prkag3 gene is located in the RN locus of chromosome 15.
CC   Mutation in Prkag3 results in an altered regulation of carbohydrate
CC   metabolism, particularly in skeletal muscle. PRKAG3 is useful as
CC   therapeutic for treating carbohydrate metabolism disorders such as
CC   diabetes, obesity, and disorders associated with muscle metabolism
CC   such as myopathy and cardiovascular diseases, to modulate AMPK
CC   activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC   and its functionally altered mutants are useful for the diagnostic
CC   evaluation, genetic testing and prognosis of a metabolic disorder,
CC   preferably a carbohydrate metabolism disorder. Primers that can detect
CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
CC   from the expression of a functionally altered allele of PRKAG3.
CC   Transgenic animal and host cell transformed with PRKAG3 or a
CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC   screening compounds able to modulate AMPK activity. Nucleic acid
CC   encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
CC   of PRKAG3 and is useful in gene therapy.
XX
SQ   Sequence 1867 BP; 380 A; 583 C; 529 G; 375 T; 0 other;


  Query Match             8.6%;  Score 148;  DB 22;  Length 1867;
  Best Local Similarity   88.9%;  Pred. No. 5.8e-29;
  Matches  160;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy   1510 ccatcctaaccagggttccctgctgccccggccctccttcctctaccgcactatccaaga 1569
           ||  |  | |   ||  |||||||||||||||||||||||||||||||||||| ||||||||
Db    909 cctgcacatctttggcaccctgctgccccggccctccttcctctaccgcaccatccaaga 968

Qy   1570 tttgggcatcggcacattccgagacttggctgtggtgctggagacagcacccatcctgac 1629
           |||||||||||||||||||||||||||||| |||||||||| || || ||||||||||
Db    969 tttgggcatcggcacattccgagacttggccgtggtgctggaaacggcgcccatcctgac 1028

Qy   1630 tgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtgg 1689
           |||||||||||||| ||||||||||||||||||| |||||||||||||||  |||
Db   1029 cgcactggacatcttcgtggaccggcgtgtgtctgcgctgcctgtggtcaacgaaactgg 1088


RESULT  15
AAD03319
ID   AAD03319 standard; cDNA; 1873 BP.
XX
AC   AAD03319;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Pig AMPK gamma subunit muscle-specific isoform, complete PRKAG3 cDNA.
XX
KW   Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW   PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW   genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW   cystathione beta synthase; CBS; cardiant; gene therapy; RN locus;
KW   chromosome 15; ss.
XX
OS   Sus scrofa.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1395
FT                   /*tag= a
FT                   /product= "Sus scrofa complete Prkag3 protein"
XX
PN   WO200120003-A2.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-EP09896.
XX
PR   10-SEP-1999;  99EP-0402236.
PR   18-MAY-2000; 2000EP-0401388.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA   (ANDE/) ANDERSSON L.
PA   (LOOF/) LOOFT C.
PA   (KALM/) KALM E.
XX
PI   Andersson L,  Looft C,  Kalm E,  Milan D,  Robic A,  Rogel-Gaillard C;
PI   Iannuccelli N,  Gellin J,  Le Roy P,  Chardon P;
XX
DR   WPI; 2001-244810/25.
DR   P-PSDB; AAE00222.
XX
PT   New variants of the gamma subunit of vertebrate adenosine
PT   monophosphate-activated kinase for diagnosis or treatment of disorders
PT   associated with energy metabolism such as diabetes, obesity, and
PT   myopathy  -
XX
PS   Claim 12; Page 62-64; 71pp; English.
XX
CC   The present sequence is a cDNA encoding pig adenosine monophosphate
CC   (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC   complete PRKAG3. Prkag3 gene is located in the RN locus of chromosome
CC   15. Mutation in Prkag3 results in an altered regulation of carbohydrate
CC   metabolism, particularly in skeletal muscle. PRKAG3 is useful as
CC   therapeutic for treating carbohydrate metabolism disorders such as
CC   diabetes, obesity, and disorders associated with muscle metabolism
CC   such as myopathy and cardiovascular diseases, to modulate AMPK
CC   activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC   and its functionally altered mutants are useful for the diagnostic
CC   evaluation, genetic testing and prognosis of a metabolic disorder,
CC   preferably a carbohydrate metabolism disorder. Primers that can detect
CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
CC   from the expression of a functionally altered allele of PRKAG3.
CC   Transgenic animal and host cell transformed with PRKAG3 or a
CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC   screening compounds able to modulate AMPK activity. Nucleic acid
CC   encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
CC   of PRKAG3 and is useful in gene therapy.
XX
SQ   Sequence 1873 BP; 382 A; 580 C; 535 G; 376 T; 0 other;


  Query Match             8.6%;  Score 148;  DB 22;  Length 1873;
  Best Local Similarity   88.9%;  Pred. No. 5.8e-29;
  Matches  160;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy   1510 ccatcctaaccagggttccctgctgccccggccctccttcctctaccgcactatccaaga 1569
           ||  |  | |   ||  |||||||||||||||||||||||||||||||||||| ||||||||
Db    915 cctgcacatctttggcaccctgctgccccggccctccttcctctaccgcaccatccaaga 974

Qy   1570 tttgggcatcggcacattccgagacttggctgtggtgctggagacagcacccatcctgac 1629
           |||||||||||||||||||||||||||||| |||||||||| || || ||||||||||
Db    975 tttgggcatcggcacattccgagacttggccgtggtgctggaaacggcgcccatcctgac 1034

Qy   1630 tgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtgg 1689
```

```
                ||||||||||||||| |||||||||||||||||||| |||||||||||||||||  |||
Db    1035 cgcactggacatcttcgtggaccggcgtgtgtctgcgctgcctgtggtcaacgaaactgg 1094
```

Search completed: October 3, 2002, 16:30:16
Job time: 14286 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 11:39:20 ; Search time 180.77 Seconds
(without alignments)
2339.884 Million cell updates/sec

Title: US-09-826-581-3
Perfect score: 1722
Sequence: 1 cctggccccctcagatcaaga.....gatgagaggctcgggctgga 1722

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	68.8	4.0	1576	3	US-09-101-146-63	Sequence 63, Appl
	2	67.6	3.9	7218	1	US-08-232-463-14	Sequence 14, Appl
	3	63.8	3.7	1435	2	US-08-878-989-14	Sequence 14, Appl
	4	63.8	3.7	1435	4	US-09-272-796-14	Sequence 14, Appl
c	5	61	3.5	7218	1	US-08-232-463-14	Sequence 14, Appl
	6	39.4	2.3	289	4	US-09-007-005-17	Sequence 17, Appl
	7	39.4	2.3	289	4	US-09-244-796-17	Sequence 17, Appl
c	8	39	2.3	53526	3	US-08-658-136-2	Sequence 2, Appli
c	9	39	2.3	53577	3	US-08-658-136-1	Sequence 1, Appli
	10	38.6	2.2	320	4	US-09-165-264-7	Sequence 7, Appli
	11	38.6	2.2	320	4	US-09-165-264-14	Sequence 14, Appl
	12	37.4	2.2	319	4	US-09-165-264-8	Sequence 8, Appli
	13	36.2	2.1	320	4	US-09-165-264-13	Sequence 13, Appl
	14	36	2.1	320	4	US-09-165-264-11	Sequence 11, Appl
c	15	36	2.1	2278	1	US-07-959-943-6	Sequence 6, Appli
c	16	36	2.1	2415	1	US-07-959-943-8	Sequence 8, Appli
c	17	35.8	2.1	12001	1	US-08-458-568A-11	Sequence 11, Appl
	18	35.4	2.1	318	4	US-09-165-264-12	Sequence 12, Appl
	19	35.2	2.0	4066	3	US-09-215-966-21	Sequence 21, Appl
	20	34.4	2.0	707	4	US-08-998-416-876	Sequence 876, App
c	21	34.4	2.0	152331	3	US-09-128-155-16	Sequence 16, Appl
	22	34.2	2.0	6306	1	US-08-466-390-3	Sequence 3, Appli
	23	34.2	2.0	6306	1	US-08-470-950-3	Sequence 3, Appli
	24	34.2	2.0	6306	1	US-08-467-781-3	Sequence 3, Appli
	25	34.2	2.0	6306	1	US-08-195-487-3	Sequence 3, Appli
	26	34.2	2.0	6306	2	US-08-483-924-3	Sequence 3, Appli
	27	34.2	2.0	6306	5	PCT-US93-06160-3	Sequence 3, Appli
c	28	34.2	2.0	7705	2	US-08-687-080-115	Sequence 115, App
c	29	34	2.0	2117	4	US-09-155-768-1	Sequence 1, Appli
c	30	34	2.0	6232	4	US-08-456-200B-11	Sequence 11, Appl
c	31	34	2.0	7175	1	US-08-455-543A-8	Sequence 8, Appli
c	32	34	2.0	7175	2	US-08-193-078B-8	Sequence 8, Appli
c	33	34	2.0	7175	2	US-08-223-305C-8	Sequence 8, Appli
c	34	34	2.0	7175	2	US-08-149-097D-8	Sequence 8, Appli
c	35	34	2.0	7175	3	US-08-949-386-8	Sequence 8, Appli
c	36	34	2.0	7175	3	US-08-450-562-8	Sequence 8, Appli
c	37	34	2.0	7175	4	US-08-984-709A-8	Sequence 8, Appli
c	38	34	2.0	7177	4	US-09-268-163-7	Sequence 7, Appli
c	39	34	2.0	7266	3	US-08-713-118-1	Sequence 1, Appli
c	40	34	2.0	7266	3	US-09-452-007-1	Sequence 1, Appli
c	41	34	2.0	7362	1	US-08-455-543A-7	Sequence 7, Appli
c	42	34	2.0	7362	2	US-08-193-078B-7	Sequence 7, Appli
c	43	34	2.0	7362	2	US-08-223-305C-7	Sequence 7, Appli
c	44	34	2.0	7362	2	US-08-149-097D-7	Sequence 7, Appli
c	45	34	2.0	7362	3	US-08-949-386-7	Sequence 7, Appli

ALIGNMENTS

```
RESULT   1
US-09-101-146-63
; Sequence 63, Application US/09101146
; Patent No. 6124125
;  GENERAL INFORMATION:
;    APPLICANT:  Dartmouth College, St. Vincents Institute of
;    APPLICANT:  Medical Research, Kemp et al.
;    TITLE OF INVENTION:  No. 6124125el AMP Activated Protein Kinase
;    NUMBER OF SEQUENCES:  64
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jane Massey Licata, Esq.
;      STREET:  66 E. Main Street
;      CITY:  Marlton
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08053
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM PC
;      OPERATING SYSTEM:  WINDOWS 95
;      SOFTWARE:  WORDPERFECT 6.0 FOR WINDOWS
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/101,146
;      FILING DATE:  October 7, 1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PN7450
;      FILING DATE:  8 JAN 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jane Massey Licata
;      REGISTRATION NUMBER:  32,257
;      REFERENCE/DOCKET NUMBER:  DC-0050
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (856) 810-1515
;      TELEFAX:  (856) 810-1454
;  INFORMATION FOR SEQ ID NO:  63:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1576
;      TYPE:  Nucleic acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    ANTI-SENSE:  No
US-09-101-146-63


  Query Match             4.0%;  Score 68.8;  DB 3;  Length 1576;
  Best Local Similarity   67.4%;  Pred. No. 4.2e-09;
  Matches   97;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy   1240 cagcctgtttgaagctgtctacaccctcatcaagaaccggatccatcgcctgcctgttct 1299
```

```
Db    501 CAGCTTGTTTGATGCTGTCTCTTCATTAATTCGCAACAAGATCCACAGGCTGCCAGTTAT 560

Qy   1300 tgacccggtgtcaggcaacgtactccacatcctcacacacaaacgcctgctcaagttcct 1359
Db    561 TGACCCAGAATCAGGCAATACTTTGTACATCCTCACCCACAAGCGCATTCTGAAGTTCCT 620

Qy   1360 gcacatctttgtaagcctgggccc 1383
Db    621 CAAATTGTTTATCACTGAGTTCCC 644


RESULT   2
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;  . CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14


  Query Match             3.9%;  Score 67.6;  DB 1;  Length 7218;
  Best Local Similarity   6.2%;  Pred. No. 1.6e-08;
  Matches   25;  Conservative  223;  Mismatches  152;  Indels    0;  Gaps    0;

Qy    906 ctgacggaagggaagctgcctgggactgcaaggtgaggcaggtgaccggctcccctggcc 965
Db   1018 CTCACAGAATTAATTCCGAGCTTGGCTGCAGGTCGAGGGAGCTTGCGATYYYYYYYYYYY 1077

Qy    966 tgactctggctctttctgcagagatctacctgcaaggctgcttcaagcctctggtctcca 1025
Db   1078 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1137

Qy   1026 tctctcctaatgataggtgggtgtctctgctcattcacctgagcctcctcctcccacagt 1085
Db   1138 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1197

Qy   1086 ccccttccccagtcccactcagctctgaactcacctcttcatcctaggcggcacacagac 1145
Db   1198 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1257

Qy   1146 aagggagccttggtgccctgccctcctttttaggggcctgggatggaggttgtctctccc 1205
Db   1258 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1317

Qy   1206 taggctgccccgaggctcactgctcccatctctgcagcctgtttgaagctgtctacaccc 1265
Db   1318 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1377

Qy   1266 tcatcaagaaccggatccatcgcctgcctgttcttgaccc 1305
Db   1378 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1417


RESULT   3
US-08-878-989-14
; Sequence 14, Application US/08878989
; Patent No. 5885803
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Guegler, Karl G.
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  DISEASE ASSOCIATED PROTEIN
;    TITLE OF INVENTION:  KINASES
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/878,989
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J J
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0321 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1435 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  PENITUT01
;      CLONE:  1452972
```

US-08-878-989-14

Query Match 3.7%; Score 63.8; DB 2; Length 1435;
Best Local Similarity 67.9%; Pred. No. 8.8e-08;
Matches 89; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
Qy   1241 agcctgtttgaagctgtctacaccctcatcaagaaccggatccatcgcctgcctgttctt 1300
          ||||| || || ||||| ||| || | ||||| ||    |||||  |  |||| ||| ||
Db    509 AGCCTCTTCGATGCTGTATACTCCTTGATCAAAAATAAAATCCACAGATTGCCCGTTATT 568

Qy   1301 gacccggtgtcaggcaacgtactccacatcctcacacacaaacgcctgctcaagttcctg 1360
          |||||  |    || || | |||  | || || || ||||||  |  | ||||||||||
Db    569 GACCCTATCAGTGGGAATGCACTTTATATACTTACCCACAAAAGAATCCTCAAGTTCCTC 628

Qy   1361 cacatctttgt 1371
          ||  | ||| |
Db    629 CAGCTTTTTAT 639
```

RESULT 4
US-09-272-796-14
; Sequence 14, Application US/09272796
; Patent No. 6207148
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
; TITLE OF INVENTION: KINASES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/272,796
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/878,989
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0321 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1435 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PENITUT01
; CLONE: 1452972
US-09-272-796-14

Query Match 3.7%; Score 63.8; DB 4; Length 1435;
Best Local Similarity 67.9%; Pred. No. 8.8e-08;
Matches 89; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

```
Qy   1241 agcctgtttgaagctgtctacaccctcatcaagaaccggatccatcgcctgcctgttctt 1300
          ||||| || || ||||| ||| || | ||||| ||    |||||  |  |||| ||| ||
Db    509 AGCCTCTTCGATGCTGTATACTCCTTGATCAAAAATAAAATCCACAGATTGCCCGTTATT 568

Qy   1301 gacccggtgtcaggcaacgtactccacatcctcacacacaaacgcctgctcaagttcctg 1360
          |||||  |    || || | |||  | || || || ||||||  |  | ||||||||||
Db    569 GACCCTATCAGTGGGAATGCACTTTATATACTTACCCACAAAAGAATCCTCAAGTTCCTC 628

Qy   1361 cacatctttgt 1371
          ||  | ||| |
Db    629 CAGCTTTTTAT 639
```

RESULT 5
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match 3.5%; Score 61; DB 1; Length 7218;
Best Local Similarity 3.6%; Pred. No. 9.6e-07;
Matches 13; Conservative 212; Mismatches 132; Indels 0; Gaps 0;

```
Qy    611 ggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagtcc 670
```

```
Db   1402 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1343

Qy    671 tcctggagcctggtgccctagaagcccacgtctttctgacttctggagtcctgtcgatgt 730

Db   1342 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1283

Qy    731 ctctaggtccagatctatgagattgaacaacataagattgagacctggaggggtgagtgg 790

Db   1282 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1223

Qy    791 ggagaggaacccggaaaggggctgttggtgatggtgggccagggcttaaggtggaggatg 850

Db   1222 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1163

Qy    851 ggcagtggggatgtcctggagtgaacaggggagggacaataggagcctcgggtgcctgac 910

Db   1162 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1103

Qy    911 ggaagggaagctgcctgggactgcaaggtgaggcaggtgaccggctcccctggcctg 967

Db   1102 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRATCGCAAGCTCCCTCGACCTG 1046


RESULT   6
US-09-007-005-17
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
;  APPLICANT: Szostak, Jack W.
;  APPLICANT: Roberts, Richard W.
;  APPLICANT: Liu, Rihe
;  TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
;  TITLE OF INVENTION: FUSIONS
;  FILE REFERENCE: 00786/350003
;  CURRENT APPLICATION NUMBER: US/09/007,005B
;  CURRENT FILING DATE: 1998-01-14
;  EARLIER APPLICATION NUMBER: 60/035,963
;  EARLIER FILING DATE: 1997-01-27
;  EARLIER APPLICATION NUMBER: 60/064,491
;  EARLIER FILING DATE: 1997-11-06
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 289
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Translation template
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(289)
;   OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17


  Query Match             2.3%;  Score 39.4;  DB 4;  Length 289;
  Best Local Similarity   2.7%;  Pred. No. 0.16;
  Matches    6;  Conservative  107;  Mismatches  110;  Indels    0;  Gaps    0;

Qy    727 atgtctctaggtccagatctatgagattgaacaacataagattgagacctggaggggtga 786

Db     23 rurarurururarcrararururarcrararurgrnrnrsrnrnrsrnrnrsrnrnrsrn 82

Qy    787 gtggggagaggaacccggaaaggggctgttggtgatggtgggccagggcttaaggtggag 846

Db     83 rnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrn 142

Qy    847 gatgggcagtggggatgtcctggagtgaacaggggagggacaataggagcctcgggtgcc 906

Db    143 rnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrn 202

Qy    907 tgacggaagggaagctgcctgggactgcaaggtgaggcaggtg 949

Db    203 rnrsrnrnrsrnrnrsrcrargrcrurgrcrgrurararcrur 245


RESULT   7
US-09-244-796-17
; Sequence 17, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
;  APPLICANT: Szostak, Jack W.
;  APPLICANT: Roberts, Richard W.
;  APPLICANT: Liu, Rihe
;  TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
;  TITLE OF INVENTION: FUSIONS
;  FILE REFERENCE: 00786/350007
;  CURRENT APPLICATION NUMBER: US/09/244,796
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/035,963
;  EARLIER FILING DATE: 1997-01-27
;  EARLIER APPLICATION NUMBER: 60/064,491
;  EARLIER FILING DATE: 1997-11-06
;  EARLIER APPLICATION NUMBER: 09/007,005
;  EARLIER FILING DATE: 1998-01-14
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 289
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Translation template
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(289)
;   OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17


  Query Match             2.3%;  Score 39.4;  DB 4;  Length 289;
  Best Local Similarity   2.7%;  Pred. No. 0.16;
  Matches    6;  Conservative  107;  Mismatches  110;  Indels    0;  Gaps    0;

Qy    727 atgtctctaggtccagatctatgagattgaacaacataagattgagacctggaggggtga 786

Db     23 rurarurururarcrararururarcrararurgrnrnrsrnrnrsrnrnrsrnrnrsrn 82

Qy    787 gtggggagaggaacccggaaaggggctgttggtgatggtgggccagggcttaaggtggag 846

Db     83 rnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrn 142

Qy    847 gatgggcagtggggatgtcctggagtgaacaggggagggacaataggagcctcgggtgcc 906

Db    143 rnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrnrnrsrn 202

Qy    907 tgacggaagggaagctgcctgggactgcaaggtgaggcaggtg 949

Db    203 rnrsrnrnrsrnrnrsrcrargrcrurgrcrgrurararcrur 245


RESULT   8
US-08-658-136-2/c
; Sequence 2, Application US/08658136
; Patent No. 6071717
;  GENERAL INFORMATION:
;    APPLICANT:  KLINGER, KATHERINE W
;    APPLICANT:  LANDES, GREGORY M
;    APPLICANT:  BURN, TIMOTHY C
;    APPLICANT:  CONNORS, TIMOTHY D
;    APPLICANT:  DACKOWSKI, WILLIAM
;    APPLICANT:  GERMINO, GREGORY
;    APPLICANT:  QIAN, FENG
;    TITLE OF INVENTION:  POLYCYSTIC KIDNEY DISEASE GENE
;    NUMBER OF SEQUENCES:  58
```

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  GENZYME CORPORATION
;      STREET:  ONE MOUNTAIN ROAD
;      CITY:  FRAMINGHAM
;      STATE:  MASSACHUSETTS
;      COUNTRY:  USA
;      ZIP:  01701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/658,136
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LASSEN, ELIZABETH
;      REGISTRATION NUMBER:  31,845
;      REFERENCE/DOCKET NUMBER:  GEN4-17.8
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  508-872-8400
;      TELEFAX:  508-872-5415
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  53526 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-658-136-2


  Query Match             2.3%;  Score 39;  DB 3;  Length 53526;
  Best Local Similarity   52.1%;  Pred. No. 1.7;
  Matches   87;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy     777 ggaggggtgagtggggagaggaacccggaaaggggctgttggtgatggtgggccagggct 836
           ||||||  ||| |||  |||||    |||   |||  |  ||     | ||   ||||
Db   34640 GGAGGGAGGAGGGGGAGGAGGAGGGGGGAGGAGGGGAGGAGGGGAGGAGGGGG 34581

Qy     837 taaggtggaggatgggcagtggggatgtcctggagtgaacaggggagggacaataggagc 896
             |  |  | ||| ||| ||  | |  |     | | ||  |||||||||  |  | |||
Db   34580 GACAGGAGGGGAGGGGGAGGAGACGGGAGGAAGGGGGAGGAGGGGAGGGGGAGGAAGAGG 34521

Qy     897 ctcgggtgcctgacggaagggaagctgcctgggactgcaaggtgagg 943
             |||     |  |   |||| |  |   || |  |  ||| ||||
Db   34520 GAGGGGGAGGAGGGGAGGGGGAGGAGGGAGGAAGGGGGAGGGGAGG 34474


RESULT  9
US-08-658-136-1/c
; Sequence 1, Application US/08658136
; Patent No. 6071717
;  GENERAL INFORMATION:
;    APPLICANT:  KLINGER, KATHERINE W
;    APPLICANT:  LANDES, GREGORY M
;    APPLICANT:  BURN, TIMOTHY C
;    APPLICANT:  CONNORS, TIMOTHY D
;    APPLICANT:  DACKOWSKI, WILLIAM
;    APPLICANT:  GERMINO, GREGORY
;    APPLICANT:  QIAN, FENG
;    TITLE OF INVENTION:  POLYCYSTIC KIDNEY DISEASE GENE
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  GENZYME CORPORATION
;      STREET:  ONE MOUNTAIN ROAD
;      CITY:  FRAMINGHAM
;      STATE:  MASSACHUSETTS
;      COUNTRY:  USA
;      ZIP:  01701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/658,136
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LASSEN, ELIZABETH
;      REGISTRATION NUMBER:  31,845
;      REFERENCE/DOCKET NUMBER:  GEN4-17.8
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  508-872-8400
;      TELEFAX:  508-872-5415
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  53577 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-658-136-1


  Query Match             2.3%;  Score 39;  DB 3;  Length 53577;
  Best Local Similarity   52.1%;  Pred. No. 1.7;
  Matches   87;  Conservative    0;  Mismatches   80;  Indels    0;  Gaps    0;

Qy     777 ggaggggtgagtggggagaggaacccggaaaggggctgttggtgatggtgggccagggct 836
           ||||||  ||| |||  |||||    |||   |||  |  ||     | ||   ||||
Db   34635 GGAGGGAGGAGGGGGAGGAGGAGGGGGGAGGAGGGGAGGAGGGAGGAGGGGG 34576

Qy     837 taaggtggaggatgggcagtggggatgtcctggagtgaacaggggagggacaataggagc 896
             |  |  | ||| ||| ||  | |  |     | | ||  |||||||||  |  | |||
Db   34575 GACAGGAGGGGAGGGGGAGGAGAGGGGAGGAAGGGGGAGGAGGGGAGGGGGAGGAAGAGG 34516

Qy     897 ctcgggtgcctgacggaagggaagctgcctgggactgcaaggtgagg 943
             |||     |  |   |||| |  |   || |  |  ||| ||||
Db   34515 GAGGGGGAGGAGGGGAGGGGGAGGAGGGGAGGAAGGGGGAGGGGAGG 34469


RESULT  10
US-09-165-264-7
; Sequence 7, Application US/09165264
; Patent No. 6197510
; GENERAL INFORMATION:
;  APPLICANT:  Vinayagamoorthy, Thuraiayah
;  TITLE OF INVENTION:  Multi-Loci Genomic Analysis
;  FILE REFERENCE:  44747
;  CURRENT APPLICATION NUMBER: US/09/165,264
;  CURRENT FILING DATE:  1998-10-01
;  NUMBER OF SEQ ID NOS:  14
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 7
;   LENGTH: 320
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:Primer sequence
US-09-165-264-7


  Query Match             2.2%;  Score 38.6;  DB 4;  Length 320;
  Best Local Similarity   50.3%;  Pred. No. 0.27;
  Matches   95;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy     777 ggaggggtgagtggggagaggaacccggaaaggggctgttggtgatggtgggccagggct 836
           || |||| | | |||| | ||     ||   ||||  |  || |  || |||   |||
Db     126 gggggggggggggggggggggggggggggggggggggggggggggggggggggggggggg 185

Qy     837 taaggtggaggatgggcagtggggatgtcctggagtgaacaggggagggacaataggagc 896
             || ||  ||  |||  | ||||  |    || | |    |||| |||      || |
Db     186 gggggggggggggggggggggggggggggggggggggggggggggggggggggggggggg 245
```

```
Qy    897 ctcgggtgcctgacggaagggaagctgcctgggactgcaaggtgaggcaggtgaccggct 956
          ||| |   |  ||  ||| | |   |||   |   || | ||  || |   ||||
Db    246 gggggggggggggggggggggggggggggggggggggggggggggggggggggggaggct 305

Qy    957 cccctggcc 965
          ||  |||
Db    306 agccatgcc 314


RESULT  11
US-09-165-264-14
; Sequence 14, Application US/09165264
; Patent No. 6197510
; GENERAL INFORMATION:
;  APPLICANT:  Vinayagamoorthy, Thuraiayah
;  TITLE OF INVENTION:  Multi-Loci Genomic Analysis
;  FILE REFERENCE:  44747
;  CURRENT APPLICATION NUMBER: US/09/165,264
;  CURRENT FILING DATE:  1998-10-01
;  NUMBER OF SEQ ID NOS:  14
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 14
;   LENGTH: 320
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:Primer sequence
US-09-165-264-14


  Query Match             2.2%;  Score 38.6;  DB 4;  Length 320;
  Best Local Similarity   48.8%;  Pred. No. 0.27;
  Matches  104;  Conservative    0;  Mismatches  109;  Indels    0;  Gaps    0;

Qy    777 ggagggggtgagtggggagaggaacccggaaaggggctgttggtgatggtgggccagggct 836
          || |||| | | |||| | ||     ||   ||||  |  || |  || |||   |||
Db    108 gggggggggggggggggggggggggggggggggggggggggggggggggggggggggggg 167

Qy    837 taaggtggaggatgggcagtggggatgtcctggagtgaacaggggagggacaataggagc 896
            || ||  ||  |||  | ||||  |    || | |    ||||  |||      || |
Db    168 gggggggggggggggggggggggggggggggggggggggggggggggggggggggggggg 227

Qy    897 ctcgggtgcctgacggaagggaagctgcctgggactgcaaggtgaggcaggtgaccggct 956
          |||  |   |  ||  |||  |  |   |||   |   || | ||  || |   ||
Db    228 gggggggggggggggggggggggggggggggggggggggggggggggggggggggggggg 287

Qy    957 cccctggcctgactctggctctttctgcagaga 989
            ||   |  | ||  |   | | |||||||
Db    288 gggggggggggggtgtgccatgtacagcagaga 320


RESULT  12
US-09-165-264-8
; Sequence 8, Application US/09165264
; Patent No. 6197510
; GENERAL INFORMATION:
;  APPLICANT:  Vinayagamoorthy, Thuraiayah
;  TITLE OF INVENTION:  Multi-Loci Genomic Analysis
;  FILE REFERENCE:  44747
;  CURRENT APPLICATION NUMBER: US/09/165,264
;  CURRENT FILING DATE:  1998-10-01
;  NUMBER OF SEQ ID NOS:  14
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 8
;   LENGTH: 319
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:Primer sequence
US-09-165-264-8


  Query Match             2.2%;  Score 37.4;  DB 4;  Length 319;
  Best Local Similarity   50.3%;  Pred. No. 0.56;
  Matches   92;  Conservative    0;  Mismatches   91;  Indels    0;  Gaps    0;

Qy    777 ggaggggtgagtggggagaggaacccggaaaggggctgttggtgatggtgggccagggct 836
          || |||| | | |||| | ||     ||   ||||  |  || |  || |||   |||
Db    130 gggggggggggggggggggggggggggggggggggggggggggggggggggggggggggg 189

Qy    837 taaggtggaggatgggcagtggggatgtcctggagtgaacaggggagggacaataggagc 896
            || ||  ||  |||  | ||||  |    || | |    ||||  |||      || |
Db    190 gggggggggggggggggggggggggggggggggggggggggggggggggggggggggggg 249

Qy    897 ctcgggtgcctgacggaagggaagctgcctgggactgcaaggtgaggcaggtgaccggct 956
          |||  |   |  ||  |||  |  |   |||   |   || | ||  || |   ||
Db    250 gggggggggggggggggggggggggggggggggggggggggggggggggggggtctgcagca 309

Qy    957 ccc 959
          |||
Db    310 ccc 312


RESULT  13
US-09-165-264-13
; Sequence 13, Application US/09165264
; Patent No. 6197510
; GENERAL INFORMATION:
;  APPLICANT:  Vinayagamoorthy, Thuraiayah
;  TITLE OF INVENTION:  Multi-Loci Genomic Analysis
;  FILE REFERENCE:  44747
;  CURRENT APPLICATION NUMBER: US/09/165,264
;  CURRENT FILING DATE:  1998-10-01
;  NUMBER OF SEQ ID NOS:  14
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 13
;   LENGTH: 320
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:Primer sequence
US-09-165-264-13


  Query Match             2.1%;  Score 36.2;  DB 4;  Length 320;
  Best Local Similarity   49.2%;  Pred. No. 1.2;
  Matches   95;  Conservative    0;  Mismatches   98;  Indels    0;  Gaps    0;

Qy    777 ggaggggtgagtggggagaggaacccggaaaggggctgttggtgatggtgggccagggct 836
          || |||| | | |||| | ||     ||   ||||  |  || |  || |||   |||
Db    127 gggggggggggggggggggggggggggggggggggggggggggggggggggggggggggg 186

Qy    837 taaggtggaggatgggcagtggggatgtcctggagtgaacaggggagggacaataggagc 896
            || ||  ||  |||  | ||||  |    || | |    ||||  |||      || |
Db    187 gggggggggggggggggggggggggggggggggggggggggggggggggggggggggggg 246

Qy    897 ctcgggtgcctgacggaagggaagctgcctgggactgcaaggtgaggcaggtgaccggct 956
          |||  |   |  ||  |||  |  |   |||   |   || | ||  || |   ||
Db    247 gggggggggggggggggggggggggggggggggggggggggggggggggggggggacgtgg 306

Qy    957 cccctggcctgac 969
          |   ||| || |
Db    307 acgagggcatgcc 319


RESULT  14
US-09-165-264-11
; Sequence 11, Application US/09165264
; Patent No. 6197510
; GENERAL INFORMATION:
;  APPLICANT:  Vinayagamoorthy, Thuraiayah
;  TITLE OF INVENTION:  Multi-Loci Genomic Analysis
;  FILE REFERENCE:  44747
;  CURRENT APPLICATION NUMBER: US/09/165,264
;  CURRENT FILING DATE:  1998-10-01
```

```
;  NUMBER OF SEQ ID NOS:  14
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 11
;    LENGTH: 320
;    TYPE: DNA
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence:Primer sequence
US-09-165-264-11


  Query Match             2.1%;  Score 36;  DB 4;  Length 320;
  Best Local Similarity   50.6%;  Pred. No. 1.3;
  Matches   87;  Conservative    0;  Mismatches   85;  Indels    0;  Gaps    0;

Qy     777 ggaggggtgagtggggagaggaacccggaaaggggctgttggtgatggtgggccagggct 836
           || |||| | | |||| | ||     ||   ||||  |  || |  || |||   |||
Db     135 gggggggggggggggggggggggggggggggggggggggggggggggggggggggggggg 194

Qy     837 taaggtggaggatgggcagtggggatgtcctggagtgaacaggggagggacaataggagc 896
           || || ||  |||  | ||||  |    || | |    |||| |||      || |
Db     195 gggggggggggggggggggggggggggggggggggggggggggggggggggggggggggg 254

Qy     897 ctcgggtgcctgacggaagggaagctgcctgggactgcaaggtgaggcaggt 948
           ||| |   |  ||  |||  | |   |||   |   || | |||| |
Db     255 gggggggggggggggggggggggggggggggggggggggggggggggggcagat 306


RESULT  15
US-07-959-943-6/c
; Sequence 6, Application US/07959943
; Patent No. 5418162
;  GENERAL INFORMATION:
;    APPLICANT:  Blakely, Randy D.
;    APPLICANT:  Fremeau Jr., Robert T.
;    APPLICANT:  Caron, Marc G.
;    TITLE OF INVENTION:  Serotonin Transporter cDNA
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Kenneth D. Sibley; Bell, Seltzer, Park
;      ADDRESSEE:  and
;      ADDRESSEE:  Gibson
;      STREET:  Post Office Drawer 31107
;      CITY:  Raleigh
;      STATE:  No. 5418162th Carolina
;      COUNTRY:  U.S.A.
;      ZIP:  27622
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/959,943
;      FILING DATE:  19921014
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sibley, Kenneth D.
;      REGISTRATION NUMBER:  31,665
;      REFERENCE/DOCKET NUMBER:  5405.38a
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  919-881-3140
;      TELEFAX:  919-881-3175
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2278 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    POSITION IN GENOME:
;      UNITS:  2278 basepairs
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  48..1868
US-07-959-943-6


  Query Match             2.1%;  Score 36;  DB 1;  Length 2278;
  Best Local Similarity   51.9%;  Pred. No. 3;
  Matches   81;  Conservative    0;  Mismatches   75;  Indels    0;  Gaps    0;

Qy     611 ggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagtcc 670
           || ||||||||    ||||  | ||  | ||    ||||| |   |||  |||| |    |
Db    1943 GGGGAGGAGTCCAGTGGGGCTTTTCAGAGCTGAGAGAGGTGGCTGGGAAGAGGTGTCCTC 1884

Qy     671 tcctggagcctggtgccctagaagcccacgtctttctgacttctggagtcctgtcgatgt 730
           |||  | |  || | | |   |  |   ||  | ||       | ||| | |||| | |||
Db    1883 TCCCAGGGTGTGTTACACAGCATTCATGCGGATGTCCCCACACGGGATTTCTGTGGGTGT 1824

Qy     731 ctctaggtccagatctatgagattgaacaacataag 766
           ||  |   |  | | | | ||  |  | ||||
Db    1823 TTCAGGAGTGATACTTTAATAATGCGCTCCTTAAG 1788


Search completed: October  3, 2002, 16:18:58
Job time: 16778 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 14:46:08 ; Search time 7316.32 Seconds
(without alignments)
288.886 Million cell updates/sec

Title: US-09-826-581-3_COPY_600_700
Perfect score: 101
Sequence: 1 aggtccccccctggtgaggag.....tggtgccctagaagcccacg 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	101	100.0	1722	6	AX281580	AX281580 Sequence
c 2	101	100.0	196554	2	AC073128	AC073128 Homo sapi
c 3	101	100.0	206854	9	AC009974	AC009974 Homo sapi
c 4	99.4	98.4	152129	2	AC027416	AC027416 Homo sapi
5	35	34.7	2539	4	U01206	U01206 Bos taurus
c 6	35	34.7	127794	9	AL138688	AL138688 Human DNA
7	33.2	32.9	692	5	AF461698	AF461698 Gallus ga
8	33.2	32.9	895	5	S69088	S69088 gMHox-homeo
9	33.2	32.9	1084	5	CHKPRX1	D13433 Gallus gall
c 10	32.6	32.3	6335	9	D86982	D86982 Human mRNA
c 11	32.6	32.3	54018	2	AC091776	AC091776 Chlamydom
c 12	32.6	32.3	65898	2	AC090435	AC090435 Chlamydom
13	32.6	32.3	161999	2	AL133352	AL133352 Homo sapi
c 14	32.6	32.3	172748	2	AC097417	AC097417 Rattus no
c 15	32.6	32.3	203200	9	AC008744	AC008744 Homo sapi
c 16	32.4	32.1	3761	9	HSM800653	AL080141 Homo sapi
c 17	32.4	32.1	4648	9	AF274863	AF274863 Homo sapi
18	32.4	32.1	5888	4	AF214521	AF214521 Sus scrof
19	32	31.7	45375	5	AF328738	AF328738 Agelaius
c 20	32	31.7	73699	2	AC094322	AC094322 Rattus no
21	31.8	31.5	104514	2	AC008713	AC008713 Homo sapi
c 22	31.8	31.5	135354	2	AC106731	AC106731 Homo sapi
23	31.8	31.5	136657	2	AL669924	AL669924 Homo sapi
c 24	31.8	31.5	221116	2	AC032044	AC032044 Homo sapi
c 25	31.6	31.3	171428	9	AC013407	AC013407 Homo sapi
c 26	31.6	31.3	175999	2	AC021154	AC021154 Homo sapi
c 27	31.6	31.3	183491	9	CNS01DU4	AL133163 Human chr
c 28	31.4	31.1	63937	2	AC024290	AC024290 Homo sapi
c 29	31.4	31.1	67600	2	AC105210	AC105210 Homo sapi
c 30	31.4	31.1	88928	2	AC106624	AC106624 Rattus no
31	31.4	31.1	157790	2	AL357509	AL357509 Homo sapi
c 32	31.4	31.1	164959	2	AC087823	AC087823 Homo sapi
33	31.4	31.1	207636	2	AC019122	AC019122 Homo sapi
c 34	31.4	31.1	215692	2	AC090977	AC090977 Mus muscu
35	31.4	31.1	215727	2	AL358795	AL358795 Homo sapi
36	31.2	30.9	145556	2	OSJN00092	AL606651 Oryza sat
37	31.2	30.9	154140	2	AC010758	AC010758 Homo sapi
38	31.2	30.9	157437	2	AC021669	AC021669 Homo sapi
c 39	31.2	30.9	167882	2	AC096852	AC096852 Sus scrof
40	31.2	30.9	174281	2	AC068338	AC068338 Homo sapi
c 41	31.2	30.9	177433	9	AC011747	AC011747 Homo sapi
c 42	31.2	30.9	186229	2	AC095022	AC095022 Sus scrof
c 43	31.2	30.9	210651	2	AC105137	AC105137 Homo sapi
c 44	31	30.7	1105	6	AX090307	AX090307 Sequence
c 45	31	30.7	1105	9	AF305687	AF305687 Homo sapi

ALIGNMENTS

RESULT 1
AX281580
LOCUS AX281580 1722 bp DNA linear PAT 03-NOV-2001
DEFINITION Sequence 3 from Patent WO0177305.
ACCESSION AX281580
VERSION AX281580.1 GI:16608831
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Andersson,L., Luthman,H. and Marklund,S.
TITLE Variants of the human amp-activated protein kinase gamma 3 subunit
JOURNAL Patent: WO 0177305-A 3 18-OCT-2001;
Arexis AB (SE)
FEATURES Location/Qualifiers
source 1. .1722
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 321 a 504 c 534 g 363 t
ORIGIN

```
  Query Match             100.0%;  Score 101;  DB 6;  Length 1722;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-16;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 aggtccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcg 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     600 AGGTCCCCCCTGGTGAGGAGTGGGCTGGGAATCTTATGGGCACCCAGAGGGGCGGGGGCG 659

Qy      61 qaggggagtcctcctggagcctggtgccctagaagcccacg 101
           |||||||||||||||||||||||||||||||||||||||||
Db     660 GAGCGGAGTCCTCCTGGAGCCTGGTGCCCTAGAAGCCCACG 700


RESULT   2
AC073128/c
LOCUS       AC073128              196554 bp    DNA     linear   HTG 21-FEB-2001
DEFINITION  Homo sapiens chromosome 2 clone RP11-647O5, WORKING DRAFT SEQUENCE,
            17 unordered pieces.
ACCESSION   AC073128
VERSION     AC073128.3  GI:13027579
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 196554)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 196554)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JUN-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Feb 21, 2001 this sequence version replaced gi:8469048.

            ----------------- Genome Center -----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: H_NH0647O05
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 98%
            Sequencing vector: plasmid; 0%
            Chemistry: Dye-primer ET; 98% of reads
            Chemistry: Dye-terminator Big Dye; 0% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 187795 bases at least Q40
            Consensus quality: 190513 bases at least Q30
            Consensus quality: 192099 bases at least Q20
            Insert size: 200000; agarose-fp
            Insert size: 194954; sum-of-contigs
            Quality coverage: 5.58 in Q20 bases; agarose-fp
            Quality coverage: 5.67 in Q20 bases; sum-of-contigs
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 17 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1157: contig of 1157 bp in length
            *     1158     1257: gap of unknown length
            *     1258     3600: contig of 2343 bp in length
            *     3601     3700: gap of unknown length
            *     3701     5103: contig of 1403 bp in length
            *     5104     5203: gap of unknown length
            *     5204     8524: contig of 3321 bp in length
            *     8525     8624: gap of unknown length
            *     8625    11856: contig of 3232 bp in length
            *    11857    11956: gap of unknown length
            *    11957    15783: contig of 3827 bp in length
            *    15784    15883: gap of unknown length
            *    15884    21906: contig of 6023 bp in length
            *    21907    22006: gap of unknown length
            *    22007    28887: contig of 6881 bp in length
            *    28888    28987: gap of unknown length
            *    28988    35255: contig of 6268 bp in length
            *    35256    35355: gap of unknown length
            *    35356    44642: contig of 9287 bp in length
            *    44643    44742: gap of unknown length
            *    44743    58275: contig of 13533 bp in length
            *    58276    58375: gap of unknown length
            *    58376    73816: contig of 15441 bp in length
            *    73817    73916: gap of unknown length
            *    73917    92140: contig of 18224 bp in length
            *    92141    92240: gap of unknown length
            *    92241   113337: contig of 21097 bp in length
            *   113338   113437: gap of unknown length
            *   113438   130325: contig of 16888 bp in length
            *   130326   130425: gap of unknown length
            *   130426   149287: contig of 18862 bp in length
            *   149288   149387: gap of unknown length
            *   149388   196554: contig of 47167 bp in length.
FEATURES             Location/Qualifiers
     source          1. .196554
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /clone="RP11-647O5"
     misc_feature    1. .1157
                     /note="assembly_name:Contig17"
     misc_feature    1258. .3600
                     /note="assembly_name:Contig18"
     misc_feature    3701. .5103
                     /note="assembly_name:Contig19
                     clone_end:T7
                     vector_side:right"
     misc_feature    5204. .8524
                     /note="assembly_name:Contig20"
     misc_feature    8625. .11856
                     /note="assembly_name:Contig21"
     misc_feature    11957. .15783
                     /note="assembly_name:Contig22"
     misc_feature    15884. .21906
                     /note="assembly_name:Contig23"
     misc_feature    22007. .28887
                     /note="assembly_name:Contig24"
     misc_feature    28988. .35255
                     /note="assembly_name:Contig25"
     misc_feature    35356. .44642
                     /note="assembly_name:Contig26"
     misc_feature    44743. .58275
                     /note="assembly_name:Contig27"
     misc_feature    58376. .73816
                     /note="assembly_name:Contig28"
     misc_feature    73917. .92140
                     /note="assembly_name:Contig29"
     misc_feature    92241. .113337
                     /note="assembly_name:Contig30"
     misc_feature    113438. .130325
                     /note="assembly_name:Contig31"
     misc_feature    130426. .149287
                     /note="assembly_name:Contig32"
     misc_feature    149388. .196554
                     /note="assembly_name:Contig33
                     clone_end:SP6
                     vector_side:right"
BASE COUNT    52296 a  46993 c  45889 g  49770 t   1606 others
ORIGIN
```

```
  Query Match             100.0%;  Score 101;  DB 2;  Length 196554;
  Best Local Similarity   100.0%;  Pred. No. 4.4e-17;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 aggtccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcg 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   60519 AGGTCCCCCCTGGTGAGGAGTGGGCTGGGAATCTTATGGGCACCCAGAGGGGCGGGGGCG 60460

Qy      61 gaggggagtcctcctggagcctggtgccctagaagcccacg 101
           |||||||||||||||||||||||||||||||||||||||||
Db   60459 GAGGGGAGTCCTCCTGGAGCCTGGTGCCCTAGAAGCCCACG 60419


RESULT   3
AC009974/c
LOCUS       AC009974               206854 bp    DNA     linear   PRI 09-JAN-2002
DEFINITION  Homo sapiens BAC clone RP11-459I19 from 2, complete sequence.
ACCESSION   AC009974
VERSION     AC009974.9  GI:16799058
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 206854)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
REFERENCE   2  (bases 1 to 206854)
  AUTHORS   Harris,A. and Cotton,M.
  TITLE     The sequence of Homo sapiens BAC clone RP11-459I19
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 206854)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 206854)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-NOV-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 206854)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JAN-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   6  (bases 1 to 206854)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JAN-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Nov 8, 2001 this sequence version replaced gi:13431203.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0459I19
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. John D.
            McPherson, Department of Genetics, Washington University, St. Louis
            MO.  For additional information about the map position of this
            sequence, see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries.  Genomics 51:1-8.  The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at the Roswell Park Cancer Institute
            (http://bacpac.med.buffalo.edu)
            VECTOR:  pBACe3.6
            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is RP11-1077K22; the clone
            sequenced to the right is RP11-647O5.  Actual start of this clone
            is at base position 1 of RP11-459I19; actual end is at base
            position 206854 of RP11-459I19.

            Data from AC079810 and AC073128 was used to finish this clone,
            AC009974.  Polymorphisms have been identified between AC073128 and
            AC009974.  A single plasmid region exists beteen 38812-38903.  Am
            unresolved tandem in the HERV SVA exists between 184390-185163.
            PCR suggests that approximately 1700 bps are missing.
FEATURES             Location/Qualifiers
     source          1. .206854
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2"
                     /clone="RP11-459I19"
                     /clone_lib="RPCI-11"
     misc_feature    1. .37
                     /note="match to EST BE314060 (NID:g9134719)"
     misc_feature    1. .37
                     /note="match to EST BI059713 (NID:g14467240)"
     misc_feature    1. .37
                     /note="match to EST BF183086 (NID:g11061273)"
     misc_feature    1. .37
                     /note="match to EST AL567345 (NID:g12920610)"
     misc_feature    1. .37
                     /note="match to EST AW880850 (NID:g8042860)"
     misc_feature    1. .37
                     /note="match to EST BF304755 (NID:g11251653)"
     misc_feature    1. .37
                     /note="similar to Homo sapiens EST BI114348
                     (NID:g14565249)"
     misc_feature    1. .37
                     /note="match to EST BG477625 (NID:g13409904)"
     misc_feature    1. .37
                     /note="match to EST BE047599 (NID:g8364652) tz39c01.y1"
     misc_feature    3. .37
                     /note="match to EST BE908408 (NID:g10402954)"
     misc_feature    4. .37
                     /note="match to EST AI670836 (NID:g4850567) wa04g10.x1"
     misc_feature    164. .662
                     /note="match to EST C05773 (NID:g1502549)"
     misc_feature    220. .221
                     /note="match to EST BG470047 (NID:g13402322)"
     misc_feature    281. .344
                     /note="match to EST BE047599 (NID:g8364652) tz39c01.y1"
     misc_feature    281. .344
                     /note="match to EST BE908408 (NID:g10402954)"
```

```
  misc_feature    281. .344
                  /note="match to EST BF304755 (NID:g11251653)"
  misc_feature    281. .344
                  /note="match to EST BF183086 (NID:g11061273)"
  misc_feature    281. .344
                  /note="similar to Homo sapiens EST BI114348
                  (NID:g14565249)"
  misc_feature    281. .344
                  /note="match to EST BG477625 (NID:g13409904)"
  misc_feature    281. .344
                  /note="match to EST BI059713 (NID:g14467240)"
  misc_feature    281. .344
                  /note="match to EST AL567345 (NID:g12920610)"
  misc_feature    281. .344
                  /note="match to EST BE314060 (NID:g9134719)"
  misc_feature    281. .344
                  /note="match to EST AI670836 (NID:g4850567) wa04g10.x1"
  misc_feature    281. .344
                  /note="match to EST AW880850 (NID:g8042860)"
  misc_feature    294. .344
                  /note="match to EST BG470047 (NID:g13402322)"
  misc_feature    complement(557. .558)
                  /note="match to EST AA043371 (NID:g1521226) zk53e10.r1"
  misc_feature    588. .929
                  /note="match to EST T64073 (NID:g667938) yc05d12.r1"
  misc_feature    594. .763
                  /note="match to EST BF304755 (NID:g11251653)"
  misc_feature    594. .763
                  /note="match to EST BG470047 (NID:g13402322)"
  misc_feature    594. .763
                  /note="match to EST AW880850 (NID:g8042860)"
  misc_feature    594. .731
                  /note="match to EST BE314060 (NID:g9134719)"
  misc_feature    594. .764
                  /note="match to EST BI059713 (NID:g14467240)"
  misc_feature    594. .763
                  /note="match to EST BG477625 (NID:g13409904)"
  misc_feature    594. .742
                  /note="match to EST BE047599 (NID:g8364652) tz39c01.y1"
  misc_feature    594. .763
                  /note="match to EST BE908408 (NID:g10402954)"
  misc_feature    594. .763
                  /note="match to EST AA043371 (NID:g1521226) zk53e10.r1"
  misc_feature    594. .763
                  /note="match to EST AI670836 (NID:g4850567) wa04g10.x1"
  misc_feature    594. .763
                  /note="match to EST BF183086 (NID:g11061273)"
  misc_feature    594. .763
                  /note="similar to Homo sapiens EST BI114348
                  (NID:g14565249)"
  misc_feature    599. .763
                  /note="match to EST AL567345 (NID:g12920610)"
  misc_feature    617. .1084
                  /note="match to EST AA481361 (NID:g2210913) zv44e01.r1"
  misc_feature    622. .763
                  /note="match to EST AI860958 (NID:g5514574) wl56f05.x1"
  misc_feature    622. .763
                  /note="similar to Mus musculus EST AI196847 (NID:g3749453)
                  ui67e01.x1"
  misc_feature    684. .763
                  /note="match to EST BG992568 (NID:g14396638)"
  misc_feature    962. .1084
                  /note="match to EST AI656812 (NID:g4740791) tt54b06.x1"
  misc_feature    967. .1084
                  /note="match to EST BE908408 (NID:g10402954)"
  misc_feature    967. .1085
                  /note="match to EST BF304755 (NID:g11251653)"
  misc_feature    967. .1091
                  /note="match to EST AA043371 (NID:g1521226) zk53e10.r1"
  misc_feature    967. .1071

Query Match             100.0%;  Score 101;  DB 9;  Length 206854;
Best Local Similarity   100.0%;  Pred. No. 4.3e-17;

  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 aggtccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcg 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 166099 AGGTCCCCCCTGGTGAGGAGTGGGCTGGGAATCTTATGGGCACCCAGAGGGGCGGGGGCG 166040

Qy      61 gaggggagtcctcctggagcctggtgccctagaagcccacg 101
           |||||||||||||||||||||||||||||||||||||||||
Db 166039 GAGGGGAGTCCTCCTGGAGCCTGGTGCCCTAGAAGCCCACG 165999


RESULT   4
AC027416/c
LOCUS       AC027416               152129 bp    DNA     linear   HTG 07-JUN-2000
DEFINITION  Homo sapiens clone RP11-504G11, WORKING DRAFT SEQUENCE, 32
            unordered pieces.
ACCESSION   AC027416
VERSION     AC027416.2  GI:8317289
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 152129)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-504G11
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 152129)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jun 7, 2000 this sequence version replaced gi:7342115.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L7458
                Center clone name: 504_G_11
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 135376 bases at least Q40
                Consensus quality: 143264 bases at least Q30
                Consensus quality: 146503 bases at least Q20
                Insert size: 161000; agarose-fp
                Insert size: 149029; sum-of-contigs
```

```
    Quality coverage: 3.1 in Q20 bases; agarose-fp
    Quality coverage: 3.3 in Q20 bases; sum-of-contigs
--------------.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 32 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1     1005: contig of 1005 bp in length
*     1006 1105: gap of      100 bp
*     1106     2402: contig of 1297 bp in length
*     2403 2502: gap of      100 bp
*     2503     3823: contig of 1321 bp in length
*     3824 3923: gap of      100 bp
*     3924     5020: contig of 1097 bp in length
*     5021 5120: gap of      100 bp
*     5121     6161: contig of 1041 bp in length
*     6162 6261: gap of      100 bp
*     6262     7547: contig of 1286 bp in length
*     7548 7647: gap of      100 bp
*     7648     9983: contig of 2336 bp in length
*     9984 10083: gap of      100 bp
*    10084    12556: contig of 2473 bp in length
*    12557 12656: gap of      100 bp
*    12657    15043: contig of 2387 bp in length
*    15044 15143: gap of      100 bp
*    15144    17123: contig of 1980 bp in length
*    17124 17223: gap of      100 bp
*    17224    19466: contig of 2243 bp in length
*    19467 19566: gap of      100 bp
*    19567    21928: contig of 2362 bp in length
*    21929 22028: gap of      100 bp
*    22029    24319: contig of 2291 bp in length
*    24320 24419: gap of      100 bp
*    24420    27059: contig of 2640 bp in length
*    27060 27159: gap of      100 bp
*    27160    30170: contig of 3011 bp in length
*    30171 30270: gap of      100 bp
*    30271    33968: contig of 3698 bp in length
*    33969 34068: gap of      100 bp
*    34069    38179: contig of 4111 bp in length
*    38180 38279: gap of      100 bp
*    38280    42366: contig of 4087 bp in length
*    42367 42466: gap of      100 bp
*    42467    46365: contig of 3899 bp in length
*    46366 46465: gap of      100 bp
*    46466    51285: contig of 4820 bp in length
*    51286 51385: gap of      100 bp
*    51386    55871: contig of 4486 bp in length
*    55872 55971: gap of      100 bp
*    55972    60595: contig of 4624 bp in length
*    60596 60695: gap of      100 bp
*    60696    66595: contig of 5900 bp in length
*    66596 66695: gap of      100 bp
*    66696    73218: contig of 6523 bp in length
*    73219 73318: gap of      100 bp
*    73319    77115: contig of 3797 bp in length
*    77116 77215: gap of      100 bp
*    77216    85022: contig of 7807 bp in length
*    85023 85122: gap of      100 bp
*    85123    93314: contig of 8192 bp in length
*    93315 93414: gap of      100 bp
*    93415   101193: contig of 7779 bp in length
*   101194 101293: gap of      100 bp
*   101294   113090: contig of 11797 bp in length
*   113091 113190: gap of      100 bp
*   113191   123496: contig of 10306 bp in length
*   123497 123596: gap of      100 bp
*   123597   137837: contig of 14241 bp in length
*   137838 137937: gap of      100 bp
*   137938   152129: contig of 14192 bp in length.
FEATURES             Location/Qualifiers
     source          1. .152129
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-504G11"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1005
                     /note="assembly_fragment"
     misc_feature    1106. .2402
                     /note="assembly_fragment"
     misc_feature    2503. .3823
                     /note="assembly_fragment"
     misc_feature    3924. .5020
                     /note="assembly_fragment"
     misc_feature    5121. .6161
                     /note="assembly_fragment"
     misc_feature    6262. .7547
                     /note="assembly_fragment"
     misc_feature    7648. .9983
                     /note="assembly_fragment"
     misc_feature    10084. .12556
                     /note="assembly_fragment"
     misc_feature    12657. .15043
                     /note="assembly_fragment"
     misc_feature    15144. .17123
                     /note="assembly_fragment"
     misc_feature    17224. .19466
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    19567. .21928
                     /note="assembly_fragment"
     misc_feature    22029. .24319
                     /note="assembly_fragment"
     misc_feature    24420. .27059
                     /note="assembly_fragment"
     misc_feature    27160. .30170
                     /note="assembly_fragment"
     misc_feature    30271. .33968
                     /note="assembly_fragment"
     misc_feature    34069. .38179
                     /note="assembly_fragment"
     misc_feature    38280. .42366
                     /note="assembly_fragment"
     misc_feature    42467. .46365
                     /note="assembly_fragment"
     misc_feature    46466. .51285
                     /note="assembly_fragment"
     misc_feature    51386. .55871
                     /note="assembly_fragment"
     misc_feature    55972. .60595
                     /note="assembly_fragment"
     misc_feature    60696. .66595
                     /note="assembly_fragment"
     misc_feature    66696. .73218
                     /note="assembly_fragment"
     misc_feature    73319. .77115
                     /note="assembly_fragment"
     misc_feature    77216. .85022
                     /note="assembly_fragment"
     misc_feature    85123. .93314
                     /note="assembly_fragment"
     misc_feature    93415. .101193
                     /note="assembly_fragment"
     misc_feature    101294. .113090
                     /note="assembly_fragment"

  Query Match             98.4%;  Score 99.4;  DB 2;  Length 152129;
  Best Local Similarity   99.0%;  Pred. No. 1.2e-16;
  Matches  100;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 aggtccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcg 60
```

```
          ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db  36019 AGGTCCCCCCTGGTGAGGAGTGGGCTGGGAATCTTATGGGCAGCCAGAGGGGCGGGGGCG 35960

Qy     61 gaggggagtcctcctggagcctggtgccctagaagcccacg 101
          |||||||||||||||||||||||||||||||||||||||||
Db  35959 GAGGGGAGTCCTCCTGGAGCCTGGTGCCCTAGAAGCCCACG 35919


RESULT   5
U01206
LOCUS       U01206                  2539 bp    mRNA    linear   MAM 14-APR-1994
DEFINITION  Bos taurus GRK5 protein kinase mRNA, complete cds.
ACCESSION   U01206
VERSION     U01206.1  GI:437105
KEYWORDS    .
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 2539)
  AUTHORS   Premont,R.T., Koch,W.J., Inglese,J. and Lefkowitz,R.J.
  TITLE     Identification, purification and characterization of GRK5, a member
            of the family of G protein-coupled receptor kinases
  JOURNAL   J. Biol. Chem. 269, 6832-6841 (1994)
  MEDLINE   94165084
REFERENCE   2  (bases 1 to 2539)
  AUTHORS   Premont,R.T.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-SEP-1993) Richard T. Premont, Department of Medicine,
            Duke University Medical Center, Durham, NC 27710, USA
FEATURES             Location/Qualifiers
     source          1. .2539
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /tissue_type="tongue epithelium"
                     /dev_stage="adult"
     mat_peptide     175. .1944
                     /evidence=experimental
                     /product="GRK5 protein kinase"
     CDS             175. .1947
                     /codon_start=1
                     /product="GRK5 protein kinase"
                     /protein_id="AAA17561.1"
                     /db_xref="GI:437106"
                     /translation="MELENIVANTVLLKAREGGGKRKGKSKKWKEILKFPHINQCED
                     LRRTIDRDYCSLCDKQPVGRLLFRQFCETRPGLESYIQFLDSVAEYEVTPDEKLGEKG
                     KEIMTKYLTPKSPVFITQVGRDLVSQTEEKLLQKPCKELFSACVQSVHDYLRGEPFHE
                     YLDSMYFDRFLQWKWLERQPVTKNTFRQYRVLGKGGFGEVCACQVRATGKMYACKRLE
                     KKRIKKRKGESMALNEKQILEKVNSRFVVNLAYAYETKDALCLVLTIMNGGDLKFHIY
                     NMGNPGFEEERALFYAAEILCGLEDLHHENIVYRDLKPENILLDDYGHIRISDLGLAV
                     KIPEGDLIRGRVGTVGYMAPEVLNNQRYGLSPDYWGLGCLIYEMIEGQSPFRGRKEKV
                     KREEVDRRVLETEEVYSHKFSEEAKSICKMLLTKDAKQRLGCQEEGAAEVKRHPFFRN
                     MNFKRLEAGMLDPPFVPDPRAVYCKDVLDIEQFSTVKGVNLDHTDDDFYSKFSTGSVP
                     IPWQSEMIETECFKELNVFGPHGTLSPDLNRSHPPEPPKKGLLQRLFKRQHQNNSKSS
                     PNSKTSFNHHINSNHVSSNSTGSS"
BASE COUNT      673 a    641 c    712 g    513 t
ORIGIN


  Query Match           34.7%;  Score 35;  DB 4;  Length 2539;
  Best Local Similarity  59.6%;  Pred. No. 15;
  Matches   59;  Conservative    0;  Mismatches   40;  Indels    0;  Gaps    0;

Qy      2 ggtccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcgg 61
          || |||  ||  | || ||  |  | | || |  | | || ||||    |||||| | |
Db   1986 GGACCCTTCTTCTTAGAAGCAGAATTGCAACCAGAGGAGCTCCCACCCTGGCGGGCGGGC 2045

Qy     62 aggggagtcctcctggagcctggtgccctagaagcccac 100
          ||| || ||||| |||||| ||       |  ||||||
Db   2046 GGGGAAGCCCTCCAGGAGCCGGGGAAGGAGGCGGCCCAC 2084
```

```
RESULT   6
AL138688/c
LOCUS       AL138688               127794 bp    DNA     linear   PRI 22-NOV-2000
DEFINITION  Human DNA sequence from clone RP11-264J4 on chromosome 13. Contains
            ESTs, STSs, GSSs and CpG islands. Contains the 3' part of the
            ZNF198 gene for zinc finger protein 198, a novel gene, a gene for a
            novel connexin (gap junction protein), a novel cyclophilin type
            peptidyl-prolyl cis-trans isomerase pseudogene and the GJB2 gene
            for gap junction protein beta 2, 26 kD (connexin 26), complete
            sequence.
ACCESSION   AL138688
VERSION     AL138688.27  GI:9944124
KEYWORDS    HTG; connexin; CpG island; cyclophilin; gap junction protein; GJB2;
            prolyl cis-trans isomerase; zinc finger; ZNF198.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 127794)
  AUTHORS   Wall,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-NOV-2000) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Aug 29, 2000 this sequence version replaced gi:9863492.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence has been finished according to sequence map criteria
            as follows. An attempt is made to resolve all sequencing problems,
            such as compressions and repeats, but not necessarily within known
            annotated human repeat sequence elements (e.g. Alu). Where the
            sequence is ambiguous, there is an annotation using the 'unsure'
            feature key.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 13, constructed by the Sanger Centre Chromosome 13
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr13
            RP11-264J4 is from the library RPCI-11.1 constructed at the Roswell
            Park Cancer Institute by the group of Pieter de Jong. For further
            details see http://bacpac.med.buffalo.edu/
            VECTOR: pBACe3.6
            IMPORTANT: This sequence is not the entire insert of clone
            RP11-264J4 It may be shorter because we sequence overlapping
            sections only once, except for a 100 base overlap.
            The true right end of clone RP11-264J4 is at 127794 in this
            sequence. The true left end of clone RP11-501K3 is at 100168 in
            this sequence. The true right end of clone RP11-110K18 is at 21477
            in this sequence.
FEATURES             Location/Qualifiers
     source          1. .127794
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="13"
                     /clone="RP11-264J4"
                     /clone_lib="RPCI-11.1"
     repeat_region   1. .69
                     /note="AluJ/FLAM repeat: matches 2. .69 of consensus"
     repeat_region   1. .70
                     /note="AluJ/FLAM repeat: matches 1. .69 of consensus"
     repeat_region   78. .371
                     /note="AluY repeat: matches 1. .294 of consensus"
     repeat_region   435. .508
                     /note="L1MB3 repeat: matches 6109. .6181 of consensus"
     repeat_region   941. .978
```

```
                     /note="19 copies 2 mer tt 78% conserved"
repeat_region        1208. .1524
                     /note="AluSx repeat: matches 6. .312 of consensus"
repeat_region        2360. .2623
                     /note="AluSx repeat: matches 43. .311 of consensus"
repeat_region        2764. .2916
                     /note="MIR repeat: matches 72. .260 of consensus"
misc_feature         3251. .3534
                     /note="match: GSS: Em:AQ053217"
repeat_region        3374. .3652
                     /note="AluYb8 repeat: matches 1. .309 of consensus"
repeat_region        4202. .4434
                     /note="MER33 repeat: matches 44. .324 of consensus
                     MER33 repeat: matches 44. .324 of consensus"
repeat_region        4573. .4877
                     /note="AluSc repeat: matches 1. .306 of consensus"
repeat_region        5887. .6025
                     /note="L1MEc repeat: matches 2194. .2318 of consensus"
repeat_region        6026. .6319
                     /note="AluJb repeat: matches 5. .298 of consensus"
repeat_region        6320. .6355
                     /note="L1MEc repeat: matches 2157. .2194 of consensus"
repeat_region        6417. .6507
                     /note="MER58 repeat: matches 176. .269 of consensus"
repeat_region        6505. .6556
                     /note="MER58 repeat: matches 5. .57 of consensus
                     MER58 repeat: matches 5. .57 of consensus"
repeat_region        6568. .6892
                     /note="L1MEc repeat: matches 2140. .2148 of consensus"
mRNA                 join(<7550. .7664,8315. .8566,9190. .9310,11356. .12268)
                     /gene="ZNF198"
                     /note="match: cDNAs: Em:AF060181 Em:AJ224901 Em:AF012126
                     Em:Y13472 Em:AB002383 Em:AF161582 Em:X95808
                     match: ESTs: Em:AW464846 Em:AA204514 Em:BE573918
                     Em:AA421981 Em:BE169588 Em:AL042736 Em:AV683458
                     Em:BE813524 Em:AA210440 Em:AW915439 Em:BE813463
                     Em:BE140140 Em:AI862769 Em:BE813467 Em:AI936478
                     Em:AW993981 Em:BE870707 Em:AW802290 Em:AW994816
                     Em:AV600551 Em:BE813473 Em:AA452383 Em:AW993938
                     Em:BE813646 Em:BE813649 Em:AW993972 Em:BE813501 Em:W36451
                     Em:BE813474 Em:AW802287 Em:AW802289 Em:BE832307
                     Em:BE813517 Em:AW993966 Em:AA161918 Em:AW995153
                     Em:AW579387 Em:AA361348 Em:AI462069 Em:D81220 Em:AI643814
                     Em:BE633738 Em:AA798296 Em:H72750 Em:AW579322 Em:AI112706
                     Em:AI642651 Em:AA851515 Em:AW556500 Em:D59995 Em:AA926013
                     Em:AW535895 Em:BE140127 Em:D60911 Em:BB530161 Em:AW107660
                     Em:AW822686 Em:AI111670 Em:AA960263 Em:AA370562 Em:C14750
                     Em:BE691037 Em:AW558100 Em:H54365 Em:BE332877 Em:BE988433
                     Em:AV613227 Em:AI481925 Em:AV260270 Em:AI613884
                     Em:BE710363 Em:N71855 Em:AI606599 Em:Aa251034 Em:AW193616
                     Em:H54287 Em:AA283078 Em:AA115047 Em:BE937769 Em:AA663837
                     Em:BE781058 Em:AA267676 Em:AV313871 Em:BE701954
                     Em:BE139491 Em:AA135989 Em:AI047555 Em:AA872477
                     Em:AA993582 Em:AA767826 Em:BE936168 Em:AW488120
                     Em:BE936169 Em:AA452155 Em:AA115537 Em:AI150524
                     Em:AI666742 Em:AI367945 Em:AA282776 Em:AI381192
                     Em:AA860937 Em:AW243913 Em:D80504 Em:AI472479 Em:AI690992
                     Em:C14354 Em:AI051311 Em:AA034499 Em:AW188335 Em:AA251580
                     Em:AA251581 Em:BE175951 Em:AA370561 Em:C14328 Em:AI273953
                     Em:AA629081 Em:C14749 Em:BE701998 Em:AW299813 Em:D11944
                     Em:AA863063 Em:BE701995 Em:AA505638 Em:AA746706 Em:N79268
                     Em:BE176093 Em:AA872945 Em:C14507 Em:BE943706 Em:AW301413
                     Em:BE701916 Em:AW885999 Em:AW902323 Em:AV613226
                     Em:AI866455 Em:AW068776 Em:AW901834 Em:AA254301"
                     /product="bA264J4.1 (zinc finger protein 198)"
                     /evidence=not_experimental
gene                 7550. .12268
                     /gene="ZNF198"
CDS                  join(<7550. .7664,8315. .8566,9190. .9310,11356. .11548)
                     /gene="ZNF198"
                     /note="continues in bA110K18 (AL137119)
                     match: proteins: Tr:O43308 Tr:Q9UBW7 Tr:O43434 Tr:O60898
                     Tr:Q9VB55 Sw:Q14202"
                     /codon_start=1
                     /evidence=not_experimental
                     /product="bA264J4.1 (zinc finger protein 198)"
                     /protein_id="CAC16956.1"
                     /db_xref="GI:11322768"
                     /translation="YLCGSNRKDNIFIDPGYQTFEQELNKILRSWQPSILPDGSIFSR
                     VEEDYLWRIKQLGSHSPVALLNTLFYFNTKYFGLKTVEQHLRLSFGTVFRHWKKNPLT
                     MENKACLRYQVSSLCGTDNEDKITTGKRKHEDDEPVFEQIENTANPSRCPVKMFECYL
                     SKSPQNLNQRMDVFYLQPECSSSTDSPVWYTSTSLDRNTLENMLVRVLLVKDIYDKDN
                     YELDEDTD"
repeat_region        8586. .8895
                     /note="AluSp repeat: matches 1. .308 of consensus"
repeat_region        9552. .9843
                     /note="AluJo repeat: matches 4. .292 of consensus"
repeat_region        9885. .10074
                     /note="MER58A repeat: matches 27. .224 of consensus"
repeat_region        10438. .10804
                     /note="MER46B repeat: matches 1. .236 of consensus
                     MER46B repeat: matches 1. .236 of consensus"
misc_feature         complement(11823. .12224)
                     /note="match: STS: Em:G38645"
polyA_signal         12205. .12210
                     /gene="ZNF198"
polyA_site           12224
                     /gene="ZNF198"
polyA_site           12225
                     /gene="ZNF198"
polyA_site           12226
                     /gene="ZNF198"
polyA_signal         12257. .12262
                     /gene="ZNF198"
polyA_site           12268
                     /gene="ZNF198"
repeat_region        12269. .12338
                     /note="5 copies 14 mer 84% conserved"
misc_feature         12662. .12957
                     /note="match: GSS: Em:AQ093575"
misc_feature         complement(15734. .16180)
                     /note="match: GSS: Em:AQ142889"
repeat_region        16151. .16285
                     /note="MER46C repeat: matches 181. .333 of consensus
                     MER46C repeat: matches 181. .333 of consensus"
repeat_region        16354. .16432
                     /note="MER46C repeat: matches 2. .81 of consensus
                     MER46C repeat: matches 2. .81 of consensus"
repeat_region        16754. .16843
                     /note="L2 repeat: matches 2641. .2750 of consensus"
misc_feature         complement(16918. .17359)
                     /note="match: STS: Em:G23181"
repeat_region        17228. .17305
                     /note="L2 repeat: matches 2668. .2746 of consensus"
repeat_region        17564. .17668
                     /note="MER5A repeat: matches 57. .161 of consensus"
repeat_region        18161. .18307
                     /note="MIR repeat: matches 6. .174 of consensus"
repeat_region        18633. .18949

  Query Match             34.7%;  Score 35;  DB 9;  Length 127794;
  Best Local Similarity   66.7%;  Pred. No. 6.3;
  Matches   50;  Conservative    0;  Mismatches   25;  Indels    0;  Gaps    0;

Qy       5 ccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggagg 64
           |||| | ||  ||| |||||  ||||| |  | |||||||| | ||||  || | ||
Db  44458 CCCCGCGGGGAAGGGGTGGGGAGGGAAGCGAAGGGGCACCCCGCGGGGGAGGCGAGGGAA 44399

Qy      65 ggagtcctcctggag 79
           || | | ||| | ||
Db  44398 GGGGGCATCCCGCAG 44384


RESULT   7
AF461698
LOCUS       AF461698                692 bp    mRNA    linear   VRT 24-JAN-2002
```

```
DEFINITION  Gallus gallus paired-related homeobox mRNA, partial cds.
ACCESSION   AF461698
VERSION     AF461698.1  GI:18308153
KEYWORDS    .
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 692)
  AUTHORS   Venkataramu,C.R. and Sokolowski,B.H.A.
  TITLE     Protein-protein interactions of Kvbeta
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 692)
  AUTHORS   Venkataramu,C.R. and Sokolowski,B.H.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-DEC-2001) Otolaryngology, University of South
            Florida, 12901 Bruce B. Downs Blvd., Tampa, FL 33612, USA
FEATURES             Location/Qualifiers
     source          1. .692
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /tissue_type="inner ear"
                     /dev_stage="embryonic"
     CDS             52. .>692
                     /codon_start=1
                     /product="paired-related homeobox"
                     /protein_id="AAL67846.1"
                     /db_xref="GI:18308154"
                     /translation="MASSYAHAMERQALLPARLDGPAGLDNLQAKKNFSVSHLLDLEE
                     AGDMVAAQGDEGGGEPGRSLLESPGLTSGSDTPQQDNDQLNSEEKKKRKQRRNRTTFN
                     SSQLQALERVFERTHYPDAFVREDLARRVNLTEARVQVWFQNRRAKFRRNERAMLASK
                     NASLLKSYSGDVTAVEQPIVPRPAPRPTDYLSWGTASPYSAMATYSTTCTNASP"
BASE COUNT      158 a    220 c    217 g     97 t
ORIGIN


  Query Match             32.9%;  Score 33.2;  DB 5;  Length 692;
  Best Local Similarity   67.1%;  Pred. No. 59;
  Matches   47;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy      17 ggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagtcctcctg 76
           ||||  ||| ||  |  |  ||| | | ||| ||| || |||||| || ||| ||  | |
Db     177 GGAGGAGGCGGGCGACATGGTGGCCGCGCAGGGGGACGAGGGCGGCGGCGAGCCCGGCCG 236

Qy      77 gagcctggtg 86
           |||| || ||
Db     237 GAGCTTGTTG 246


RESULT   8
S69088
LOCUS       S69088                  895 bp    mRNA    linear   VRT 23-SEP-1994
DEFINITION  gMHox=homeobox [chickens, stage 17 embryo, heart, mRNA, 895 nt].
ACCESSION   S69088
VERSION     S69088.1  GI:545359
KEYWORDS    .
SOURCE      chicken heart stage 17 embryo.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 895)
  AUTHORS   Kuratani,S., Martin,J.F., Wawersik,S., Lilly,B., Eichele,G. and
            Olson,E.N.
  TITLE     The expression pattern of the chick homeobox gene gMHox suggests a
            role in patterning of the limbs and face and in
            compartmentalization of somites
  JOURNAL   Dev. Biol. 161 (2), 357-369 (1994)
  MEDLINE   94148118
  REMARK    GenBank staff at the National Library of Medicine created this
            entry [NCBI gibbsq 144347] from the original journal article.
            This sequence comes from Fig. 1A.
FEATURES             Location/Qualifiers
     source          1. .895
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
     gene            1. .895
                     /gene="gMHox"
                     /note="homeobox"
     CDS             31. .768
                     /gene="gMHox"
                     /note="homeodomain protein; This sequence comes from Fig.
                     1A"
                     /codon_start=1
                     /protein_id="AAB29880.1"
                     /db_xref="GI:545360"
                     /translation="MASSYAHAMERQALLPARLDGPAGLDNLQAKKNFSVSHLLDLEE
                     AGDMVAAQGDEGGGEPGRSLLESPGLTSGSDTPQQDNDQLNSEEKKKRKQRRNRTTFN
                     SSQLQALERVFERTHYPDAFVREDLARRVNLTEARVQVWFQNRRAKFRRNERAMLASK
                     NASLLKSYSGDVTAVEQPIVPRPAPRPTDYLSWGTASPYSAMATYSTTCTNASPAQGM
                     NMANSIANLRLKAKEYSLQRNQVPTVN"
BASE COUNT      254 a    255 c    246 g    140 t
ORIGIN


  Query Match             32.9%;  Score 33.2;  DB 5;  Length 895;
  Best Local Similarity   67.1%;  Pred. No. 55;
  Matches   47;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy      17 ggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagtcctcctg 76
           ||||  ||| ||  |  |  ||| | | ||| ||| || |||||| || ||| ||  | |
Db     156 GGAGGAGGCGGGCGACATGGTGGCCGCGCAGGGGGACGAGGGCGGCGGCGAGCCCGGCCG 215

Qy      77 gagcctggtg 86
           |||| || ||
Db     216 GAGCTTGTTG 225


RESULT   9
CHKPRX1
LOCUS       CHKPRX1                 1084 bp    mRNA    linear   VRT 03-FEB-1999
DEFINITION  Gallus gallus paired-related homeobox mRNA, complete cds.
ACCESSION   D13433
VERSION     D13433.1  GI:222850
KEYWORDS    paired-related homeobox.
SOURCE      Gallus gallus (library: lambda gt10) stage 24-26 cDNA to mRNA,
            clones p2 and p7.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 1084)
  AUTHORS   Nohno,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (20-OCT-1992) Tsutomu Nohno, Kawasaki Medical School,
            Dept. of Molecular Biology; 577 Matsushima, Kurashiki, Okayama
            701-01, Japan (Tel:0864-62-1111(ex.3637), Fax:0864-62-1199)
REFERENCE   2  (bases 1 to 1084)
  AUTHORS   Nohno,T., Koyama,E., Myokai,F., Taniguchi,S., Ohuchi,H., Saito,T.
            and Noji,S.
  TITLE     A chicken homeobox gene related to Drosophila paired is
            predominantly expressed in the developing limb
  JOURNAL   Dev. Biol. 158 (1), 254-264 (1993)
  MEDLINE   93321789
COMMENT     Submitted (20-OCT-1992) to DDBJ by:
            Tsutomu Nohno
            Department of Pharmacology
            Kawasaki Medical School
            577 Matsushima
            Kurashiki 701-01
            Japan
            Phone:  086-462-1111
            Fax:    086-462-1199.
FEATURES             Location/Qualifiers
     source          1. .1084
```

```
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /clone_lib="lambda gt10"
                     /dev_stage="Stage 24-26"
     gene            233. .967
                     /gene="Prx-1"
     CDS             233. .967
                     /gene="Prx-1"
                     /codon_start=1
                     /product="paired-related homeotic gene product"
                     /protein_id="BAA02695.1"
                     /db_xref="GI:222851"
                     /translation="MASSYAHAMERQALLPARLDGPAGLDNLQAKKNFSVSHLLDLEE
                     AGDMVAAQGDEGGGEPGRSLLESPGLTSGSDTPQQDNDQLNSEEKKKRKQRRNRTTFN
                     SSQLQALERVFERTHYPDAFVREDLARRVNLTEARVQVWFQNRRAKFRRNERMLASKN
                     ASLLKSYSGDVTAVEQPIVPRPAPRPTDYLSWGTASPYSAMATYSTTCTNASPAQGMN
                     MANSIANLRLKAKEYSLQRNQVPTVN"
     misc_feature    509. .691
                     /gene="Prx-1"
                     /note="homeobox"
BASE COUNT      258 a    337 c    317 g    172 t
ORIGIN
```

```
  Query Match             32.9%;  Score 33.2;  DB 5;  Length 1084;
  Best Local Similarity   67.1%;  Pred. No. 53;
  Matches   47;  Conservative    0;  Mismatches   23;  Indels    0;  Gaps    0;

Qy      17 ggagtgggctgggaatcttatgggcacccagaggggcgggggcggagggagtcctcctg 76
           ||||  ||| ||  |  |  ||| | | ||| ||| || |||||| || ||| ||  | |
Db     358 GGAGGAGGCGGGCGACATGGTGGCCGCGCAGGGGGACGAGGGCGGCGGCGAGCCCGGCCG 417

Qy      77 gagcctggtg 86
           |||| || ||
Db     418 GAGCTTGTTG 427
```

```
RESULT  10
D86982/c
LOCUS       D86982                  6335 bp    mRNA    linear   PRI 06-OCT-2001
DEFINITION  Human mRNA for KIAA0229 gene, partial cds.
ACCESSION   D86982
VERSION     D86982.1  GI:1504037
KEYWORDS    KIAA0229.
SOURCE      Homo sapiens male bone marrow myeloblast cell_line:KG-1 cDNA to
            mRNA, clone:HA2570.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 6335)
  AUTHORS   Nagase,T., Seki,N., Ishikawa,K., Ohira,M., Kawarabayasi,Y.,
            Ohara,O., Tanaka,A., Kotani,H., Miyajima,N. and Nomura,N.
  TITLE     Prediction of the coding sequences of unidentified human genes. VI.
            The coding sequences of 80 new genes (KIAA0201-KIAA0280) deduced by
            analysis of cDNA clones from cell line KG-1 and brain
  JOURNAL   DNA Res. 3 (5), 321-329 (1996)
  MEDLINE   97191544
REFERENCE   2  (bases 1 to 6335)
  AUTHORS   Ohara,O., Nagase,T., Kikuno,R. and Nomura,N.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-AUG-1996) Osamu Ohara, Kazusa DNA Research Institute;
            1532-3, Yana, Kisarazu, Chiba 292-0812, Japan
            (E-mail:cdnainfo@kazusa.or.jp, Tel:+81-438-52-3913)
FEATURES             Location/Qualifiers
     source          1. .6335
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="chromosome 6. (RH_ID :RH25435)"
                     /clone="HA2570"
                     /sex="male"
                     /cell_line="KG-1"
                     /cell_type="myeloblast"
                     /tissue_type="bone marrow"
     gene            1. .3544
                     /gene="KIAA0229"
     CDS             <1. .3544
                     /gene="KIAA0229"
                     /note="similar to human ankyrin 1(S08275)"
                     /codon_start=2
                     /protein_id="BAA13218.1"
                     /db_xref="GI:1504038"
                     /translation="SWGKGREGVVSPAGLGGALPGDGKFGSPSRLGCSLGEGVQRVAA
                     LGMGKEQELLEAARTGHLPAVEKLLSGKRLSSGFGGGGGGSGGGGGGSGGGGGGLGS
                     SSHPLSSLLSMWRGPNVNCVDSTGYTPLHHAALNGHKDVVEVLLRNDALTNVADSKGC
                     YPLHLAAWKGDAQIVRLLIHQGPSHTRVNEQNNDNETALHCAAQYGHTEVVKVLLEEL
                     TDPTMRNNKFETPLDLAALYGRLEVVKMLLNAHPNLLSCNTKKHTPLHLAARNGHKAV
                     VQVLLDAGMDSNYQTEMGSALHEAALFGKTDVVQILLAAGTDVNIKDNHGLTALDTVR
                     ELPSQKSQQIAALIEDHMTGKRSTKEVDKTPPPQPPLISSMDSISQKSQGDVEKAVTE
                     LIIDFDANAEEEGPYEALYNAISCHSLDSMASGRSSDQDSTNKEAEAAGVKPAGVRPR
                     ERPPPPAKPPPDEEEEDHIDKKYFPLTASEVLSMRPRIHGSAAREEDEHPYELLLTAE
                     TKKVVLVDGKTKDHRRSSSSRSQDSAEGQDGQVPEQFSGLLHGSSPVCEVGQDPFQLL
                     CTAGQSHPDGSPQQGACHKASMQLEETGVHAPGASQPSALDQSKRVGYLTGLPTTNSR
                     SHPETLTHTASPHPGGAEEGDRSGARSRAPPTSKPKAELKLSRSLSKSDSDLLTCSPT
                     EDATMGSRSESLSNCSIGKKRLEKSPSFASEWDEIEKIMSSIGEGIDFSQERQKISGS
                     RTLEQSVGEWLESIGLQQYESKLLLNGFDDVHFLGSNVMEEQDLRDIGISDPQHRRKL
                     LQAARSLPKVKALGYDGNSPPSVPSWLDSLGLQDYVHSFLSSGYSSIDTVKNLWELEL
                     VNVLKVQLLGHRKRIIASLADRPYEEPPQKPPRFSQLRCQDLLSQTSSPLSQNDSCTG
                     RSADLLLPPGDTGRRRHDSLHDPAAPSRAERFRIQEEHREAKLTLRPPSLAAPYAPVQ
                     SWQHQPEKLIFESCGYEANYLGSMLIKDLRGTESTQDACAKMRKSTEHMKKIPTIILS
                     ITYKGVKFIDASNKNVIAEHEIRNISCAAQDPEDLCTFAYITKDLQTSHHYCHVFSTV
                     DVNLTYEIILTLGQAFEVAYQLALQAQKSRATGASAAEMIETKSSKPVPKPRVGVRKS
                     ALEPPPDMDQDAQSHASVSWVVDPKPDSKRSLSTN"
     3'UTR           3545. .6335
BASE COUNT     1458 a   1727 c   1766 g   1384 t
ORIGIN
```

```
  Query Match             32.3%;  Score 32.6;  DB 9;  Length 6335;
  Best Local Similarity   60.9%;  Pred. No. 51;
  Matches   53;  Conservative    0;  Mismatches   34;  Indels    0;  Gaps    0;

Qy       4 tccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcggggggcggag 63
           ||  | |||| | | || |||||| | ||| |||| | |  |  || || | |
Db    2918 TCTGCACTGGGGCGTAGGGGGCTGCCAGGCTTGGGGGCCGCAGGGTCAGCTTGGCCTCAC 2859

Qy      64 gggagtcctcctggagcctggtgccct 90
           ||   |||||||||| |||| || ||
Db    2858 GGTGCTCCTCCTGGATCCTGAAGCGCT 2832
```

```
RESULT  11
AC091776/c
LOCUS       AC091776                54018 bp    DNA     linear   HTG 29-JAN-2002
DEFINITION  Chlamydomonas reinhardtii clone cr-4i21, WORKING DRAFT SEQUENCE, 5
            unordered pieces.
ACCESSION   AC091776
VERSION     AC091776.10  GI:18390210
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Chlamydomonas reinhardtii.
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 54018)
  AUTHORS   Jia,H., Lin,S., Wu,H., Dutcher,S. and Roe,B.A.
  TITLE     Chlamydomonas reinhardtii BAC Clone cr-4i21
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 54018)
  AUTHORS   Jia,H., Lin,S., Wu,H., Dutcher,S. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (02-JUN-2001) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
COMMENT     On Jan 29, 2002 this sequence version replaced gi:16271941.
            -------------- Genome Center
            Center: Department Of Chemistry And Biochemistry
            The University Of Oklahoma
```

```
            Center code:UOKNOR
            ----------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 5 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     5017: contig of 5017 bp in length
            *     5018     5117: gap of unknown length
            *     5118    10018: contig of 4901 bp in length
            *    10019    10118: gap of unknown length
            *    10119    21786: contig of 11668 bp in length
            *    21787    21886: gap of unknown length
            *    21887    33146: contig of 11260 bp in length
            *    33147    33246: gap of unknown length
            *    33247    54018: contig of 20772 bp in length.
FEATURES              Location/Qualifiers
     source           1. .54018
                      /organism="Chlamydomonas reinhardtii"
                      /db_xref="taxon:3055"
                      /clone="cr-4i21"
                      /clone_lib="cr-4i21"
BASE COUNT     9093 a  18192 c  16851 g   9482 t    400 others
ORIGIN


  Query Match             32.3%;  Score 32.6;  DB 2;  Length 54018;
  Best Local Similarity   63.3%;  Pred. No. 32;
  Matches   50;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy       3 gtccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcgga 62
           ||  |  ||| |||||   | |||| |  |   ||||||||||| | | ||||| |||||
Db   53405 GTTGCTGCTGTTGAGGGCGGTGCTGCGGCTTCCATGGGCACCCTGGGCGGCGGCGGCGGC 53346

Qy      63 ggggagtcctcctggagcc 81
           || |   |  ||  | |||
Db   53345 GGTGGCCCGGCCCTGGGCC 53327


RESULT  12
AC090435/c
LOCUS       AC090435               65898 bp    DNA     linear   HTG 29-JAN-2002
DEFINITION  Chlamydomonas reinhardtii clone cr-32m22, WORKING DRAFT SEQUENCE, 4
            unordered pieces.
ACCESSION   AC090435
VERSION     AC090435.16  GI:18390207
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Chlamydomonas reinhardtii.
  ORGANISM  Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE   1  (bases 1 to 65898)
  AUTHORS   Wu,H., Lin,S., Jia,H., Dutcher,S. and Roe,B.A.
  TITLE     Chlamydomonas reinhardtii BAC Clone cr-32m22
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 65898)
  AUTHORS   Wu,H., Lin,S., Jia,H., Dutcher,S. and Roe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (21-FEB-2001) Department Of Chemistry And Biochemistry,
            The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
            OK 73019, USA
COMMENT     On Jan 29, 2002 this sequence version replaced gi:18139395.
            ---------- Genome Center
            Center: Department Of Chemistry And Biochemistry
            The University Of Oklahoma
            Center code:UOKNOR
            ----------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 4 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     4454: contig of 4454 bp in length
            *     4455     4554: gap of unknown length
            *     4555    18724: contig of 14170 bp in length
            *    18725    18824: gap of unknown length
            *    18825    30657: contig of 11833 bp in length
            *    30658    30757: gap of unknown length
            *    30758    65898: contig of 35141 bp in length.
FEATURES              Location/Qualifiers
     source           1. .65898
                      /organism="Chlamydomonas reinhardtii"
                      /db_xref="taxon:3055"
                      /clone="cr-32m22"
                      /clone_lib="cr-32m22"
BASE COUNT     10668 a  22600 c  21674 g  10642 t    314 others
ORIGIN


  Query Match             32.3%;  Score 32.6;  DB 2;  Length 65898;
  Best Local Similarity   63.3%;  Pred. No. 30;
  Matches   50;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy       3 gtccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcgga 62
           ||  |  ||| |||||   | |||| |  |   ||||||||||| | | ||||| |||||
Db   30258 GTTGCTGCTGTTGAGGGCGGTGCTGCGGCTTCCATGGGCACCCTGGGCGGCGGCGGCGGC 30199

Qy      63 ggggagtcctcctggagcc 81
           || |   |  ||  | |||
Db   30198 GGTGGCCCGGCCCTGGGCC 30180


RESULT  13
AL133352
LOCUS       AL133352               161999 bp    DNA     linear   HTG 04-APR-2000
DEFINITION  Homo sapiens chromosome 10 clone RP11-411B6, *** SEQUENCING IN
            PROGRESS ***, 3 unordered pieces.
ACCESSION   AL133352
VERSION     AL133352.10  GI:7329506
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 161999)
  AUTHORS   Pearce,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-APR-2000) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Mar 26, 2000 this sequence version replaced gi:7327699.
            IMPORTANT: This sequence is unfinished and does not necessarily
            represent the correct sequence.  Work on the sequence is in
            progress and the release of this data is based on the understanding
            that the sequence may change as work continues.  The sequence may
            be contaminated with foreign sequence from E.coli, yeast, vector,
            phage etc. Order of segments is not known; 800 n's separate
            segments. Contig_ID: 00484  Length: 157824bp
            Contig_ID: 00808  Length: 1515bp
            Contig_ID: 01411  Length: 1060bp.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1   157824: contig of 157824 bp in length
            *   157825 158624: gap of      800 bp
```

```
            *   158625   160139: contig of 1515 bp in length
            *   160140 160939: gap of      800 bp
            *   160940   161999: contig of 1060 bp in length.
FEATURES             Location/Qualifiers
     source          1. .161999
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="10"
                     /clone="RP11-411B6"
                     /clone_lib="RPCI-11.2"
BASE COUNT    44214 a  35796 c  36927 g  43462 t   1600 others
ORIGIN


  Query Match             32.3%;  Score 32.6;  DB 2;  Length 161999;
  Best Local Similarity   60.9%;  Pred. No. 25;
  Matches   53;  Conservative    0;  Mismatches   34;  Indels    0;  Gaps    0;

Qy      10 ctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagt 69
           ||| ||| || | ||| |||       || |   ||||| |||| |||| |  ||||||
Db  39337 CTGATGAAGACTAGGCGGGGGCAAGGCTGTGGCACCAGATGGGCAGGGGTGCTGGGGAGG 39396

Qy      70 cctcctggagcctggtgccctagaagc 96
           ||    | | || | | ||||  || |
Db  39397 CCAAAAGTAACCAGCTTCCCTCCAAAC 39423


RESULT  14
AC097417/c
LOCUS       AC097417               172748 bp    DNA     linear   HTG 20-DEC-2001
DEFINITION  Rattus norvegicus clone CH230-201C5, *** SEQUENCING IN PROGRESS
            ***, 69 unordered pieces.
ACCESSION   AC097417
VERSION     AC097417.3  GI:17973606
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 172748)
  AUTHORS   Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
            Alsbrooks,S.L., Amaratunge,H.C., Are,J.R., Banks,T., Barbaria,J.,
            Benton,J., Bimage,K., Blankenburg,K., Bonnin,D., Bouck,J.,
            Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
            Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F.,
            Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
            Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
            Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C.,
            Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
            Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
            Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
            Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J.,
            Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
            Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,
            Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
            Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
            Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
            Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
            Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
            Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
            Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
            Loulseged,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
            Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
            Martinez,E., Massey,E., Mawhiney,E., McLeod,M.P., Meador,M.,
            Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
            Morgan,M., Morris,S., Moser,M., Neal,D., Newtson,J., Newtson,N.,
            Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwokenkwo,S.,
            Oguh,M., Okwuonu,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B.,
            Peery,J., Perez,L., Peters,L., Pickens,R., Primus,E., Pu,L.L.,
            Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
            Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshtari,N.,
            Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
            Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K.,
            Tang,H., Tansey,J., Taylor,C., Taylor,T., Telfrod,B., Thomas,N.,
            Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R.,
            Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
            Watlington,S., Williams,G., Williamson,A., Wleczyk,R., Wooden,S.,
            Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
            Weinstock,G. and Gibbs,R.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 172748)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-OCT-2001) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On Dec 20, 2001 this sequence version replaced gi:17064486.
            ---------- Genome Center
                Center: Baylor College of Medicine
                Center code: BCM
                Web site: http://www.hgsc.bcm.tmc.edu/
                Contact: hgsc-help@bcm.tmc.edu
            ---------- Project Information
                Center project name: GIDP
                Center clone name: CH230-201C5
            ---------- Summary Statistics
                Assembly program: Phrap; version 0.990329First call to
            findPhrapList
                Consensus quality: 140460 bases at least Q40
                Consensus quality: 148273 bases at least Q30
                Consensus quality: 155031 bases at least Q20
                Estimated insert size: 140848; sum-of-contigs estimation
                Quality coverage: 0x in Q20 bases; agarose-fp estimation
                Quality coverage: 2.1x in Q20 bases; sum-of-contigs estimation
            ----------
            * NOTE: Estimated insert size may differ from sequence length
            *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 69 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    11809: contig of 11809 bp in length
            *    11810    11909: gap of unknown length
            *    11910    17758: contig of 5849 bp in length
            *    17759    17858: gap of unknown length
            *    17859    22239: contig of 4381 bp in length
            *    22240    22339: gap of unknown length
            *    22340    25707: contig of 3368 bp in length
            *    25708    25807: gap of unknown length
            *    25808    29719: contig of 3912 bp in length
            *    29720    29819: gap of unknown length
            *    29820    33815: contig of 3996 bp in length
            *    33816    33915: gap of unknown length
            *    33916    38940: contig of 5025 bp in length
            *    38941    39040: gap of unknown length
            *    39041    43109: contig of 4069 bp in length
            *    43110    43209: gap of unknown length
            *    43210    46310: contig of 3101 bp in length
            *    46311    46410: gap of unknown length
            *    46411    50773: contig of 4363 bp in length
            *    50774    50873: gap of unknown length
            *    50874    56357: contig of 5484 bp in length
            *    56358    56457: gap of unknown length
            *    56458    59571: contig of 3114 bp in length
            *    59572    59671: gap of unknown length
            *    59672    63752: contig of 4081 bp in length
            *    63753    63852: gap of unknown length
            *    63853    67988: contig of 4136 bp in length
            *    67989    68088: gap of unknown length
            *    68089    72055: contig of 3967 bp in length
```

```
*    72056    72155: gap of unknown length
*    72156    75560: contig of 3405 bp in length
*    75561    75660: gap of unknown length
*    75661    79220: contig of 3560 bp in length
*    79221    79320: gap of unknown length
*    79321    81893: contig of 2573 bp in length
*    81894    81993: gap of unknown length
*    81994    84523: contig of 2530 bp in length
*    84524    84623: gap of unknown length
*    84624    87701: contig of 3078 bp in length
*    87702    87801: gap of unknown length
*    87802    89855: contig of 2054 bp in length
*    89856    89955: gap of unknown length
*    89956    92884: contig of 2929 bp in length
*    92885    92984: gap of unknown length
*    92985    94489: contig of 1505 bp in length
*    94490    94589: gap of unknown length
*    94590    97221: contig of 2632 bp in length
*    97222    97321: gap of unknown length
*    97322    99641: contig of 2320 bp in length
*    99642    99741: gap of unknown length
*    99742   101792: contig of 2051 bp in length
*   101793   101892: gap of unknown length
*   101893   104245: contig of 2353 bp in length
*   104246   104345: gap of unknown length
*   104346   106914: contig of 2569 bp in length
*   106915   107014: gap of unknown length
*   107015   108995: contig of 1981 bp in length
*   108996   109095: gap of unknown length
*   109096   110468: contig of 1373 bp in length
*   110469   110568: gap of unknown length
*   110569   113588: contig of 3020 bp in length
*   113589   113688: gap of unknown length
*   113689   116211: contig of 2523 bp in length
*   116212   116311: gap of unknown length
*   116312   118673: contig of 2362 bp in length
*   118674   118773: gap of unknown length
*   118774   120028: contig of 1255 bp in length
*   120029   120128: gap of unknown length
*   120129   122287: contig of 2159 bp in length
*   122288   122387: gap of unknown length
*   122388   125021: contig of 2634 bp in length
*   125022   125121: gap of unknown length
*   125122   126584: contig of 1463 bp in length
*   126585   126684: gap of unknown length
*   126685   128864: contig of 2180 bp in length
*   128865   128964: gap of unknown length
*   128965   130025: contig of 1061 bp in length
*   130026   130125: gap of unknown length
*   130126   131491: contig of 1366 bp in length
*   131492   131591: gap of unknown length
*   131592   133067: contig of 1476 bp in length
*   133068   133167: gap of unknown length
*   133168   134803: contig of 1636 bp in length
*   134804   134903: gap of unknown length
*   134904   136204: contig of 1301 bp in length
*   136205   136304: gap of unknown length
*   136305   137709: contig of 1405 bp in length
*   137710   137809: gap of unknown length
*   137810   139029: contig of 1220 bp in length
*   139030   139129: gap of unknown length
*   139130   140283: contig of 1154 bp in length
*   140284   140383: gap of unknown length
*   140384   141708: contig of 1325 bp in length
*   141709   141808: gap of unknown length
*   141809   142979: contig of 1171 bp in length
*   142980   143079: gap of unknown length
*   143080   144210: contig of 1131 bp in length
*   144211   144310: gap of unknown length
*   144311   145426: contig of 1116 bp in length
*   145427   145526: gap of unknown length
*   145527   147619: contig of 2093 bp in length
*   147620   147719: gap of unknown length
*   147720   149094: contig of 1375 bp in length
*   149095   149194: gap of unknown length
*   149195   151236: contig of 2042 bp in length
*   151237   151336: gap of unknown length
*   151337   152602: contig of 1266 bp in length
*   152603   152702: gap of unknown length
*   152703   154029: contig of 1327 bp in length
*   154030   154129: gap of unknown length
*   154130   155662: contig of 1533 bp in length
*   155663   155762: gap of unknown length

  Query Match             32.3%;  Score 32.6;  DB 2;  Length 172748;
  Best Local Similarity   58.9%;  Pred. No. 25;
  Matches   56;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy       4 tccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggag 63
           ||| ||  |||||||   | | | | | |  |   | |  || | || | || || ||||
Db  158714 TCCGCCGAGGTGAGGGCAGCGGTAGTAGTGGTGGTGACGTCCGGCGGTGGGGTGGGGGAG 158655

Qy      64 gggagtcctcctggagcctggtgccctagaagccc 98
           | | ||| || | | ||| || ||     |||| |
Db  158654 GCGCGTCTTCTTAGGGCCAGGAGCATACAAAGCTC 158620


RESULT  15
AC008744/c
LOCUS       AC008744               203200 bp    DNA     linear   PRI 22-JAN-2001
DEFINITION  Homo sapiens chromosome 19 clone CTD-2561O20, complete sequence.
ACCESSION   AC008744
VERSION     AC008744.6  GI:12331466
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 203200)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 203200)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE   3  (bases 1 to 203200)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-JAN-2001) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jan 22, 2001 this sequence version replaced gi:11178043.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 99.5% of Sequence;
            Estimated Total Number of Errors is 0.8.
            Note:
            Clipped left at overlap with AC008738.
            Clipped right at overlap with AC010485.
            STS Content:
            SHGC-35371 G28590
            WI-6761 G06103
            WI-7774 G06697
            SHGC-33750 G29472.
FEATURES             Location/Qualifiers
     source          1. .203200
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="CTD-2561O20"
BASE COUNT    52932 a  51214 c  50304 g  48750 t
ORIGIN
```

Query Match 32.3%; Score 32.6; DB 9; Length 203200;
Best Local Similarity 58.9%; Pred. No. 24;
Matches 56; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

```
Qy       7 cccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggagggg 66
           ||||||  | |||||||| ||||  |||   | ||  || |   ||| ||  |   |
Db 133557 CCCCTGCAGGGGAGTGGGGTGGGTCTCTCGAGTGCGGCCTGCCAGGCAGGAACTTGGCTC 133498

Qy      67 agtcctcctggaqcctggtgccctagaagcccacg 101
           |   | | || ||||  |   ||| | |||||| |
Db 133497 ATAGCCCATGCAGCCATGCAGCCTTGCAGCCCAGG 133463
```

Search completed: October 3, 2002, 14:46:57
Job time: 11937 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 16:30:16 ; Search time 719.93 Seconds
(without alignments)
240.868 Million cell updates/sec

Title: US-09-826-581-3_COPY_600_700
Perfect score: 101
Sequence: 1 aggtccccccctggtgaggag.....tggtgccctagaagcccacg 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	101	100.0	1722	22	AAH43683	PRKAG3 intron 4 -
	2	32.6	32.3	550	22	ABA61253	Human foetal liver
	3	32.6	32.3	550	22	ABA29092	Probe #7558 for ge
	4	32.6	32.3	550	22	AAK09549	Human brain expres
	5	32.6	32.3	550	22	AAK35441	Human bone marrow
	6	32.6	32.3	550	22	AAI17045	Probe #6978 for ge
	7	32.6	32.3	550	22	AAI41154	Probe #9840 used t
	8	32.6	32.3	680	22	AAK76637	Human immune/haema
c	9	32.6	32.3	3544	23	AAS81565	DNA encoding novel
c	10	32.4	32.1	3750	23	AAS74057	DNA encoding novel
c	11	32.4	32.1	3889	22	AAH90107	Human bone marrow
c	12	32.4	32.1	3892	22	AAH90056	Human bone marrow
c	13	32.4	32.1	4886	22	AAH89943	Human bone marrow
	14	32.2	31.9	231	22	ABA73753	Human foetal liver
	15	32.2	31.9	231	22	ABA38943	Probe #17409 for g
	16	32.2	31.9	231	22	AAK22202	Human brain expres
	17	32.2	31.9	231	22	AAK48368	Human bone marrow
	18	32.2	31.9	231	22	AAI26223	Probe #16156 for g
	19	32.2	31.9	231	22	AAI54194	Probe #22880 used
	20	31.4	31.1	17761	22	AAK79415	Human immune/haema
c	21	31	30.7	686	22	AAH08414	Human cDNA clone (
c	22	31	30.7	1030	21	AAC77940	Human cancer assoc
c	23	31	30.7	1105	22	AAF77880	Human ATFx coding
c	24	31	30.7	1374	22	AAH18673	Human cDNA sequenc
	25	30.6	30.3	412	22	AAK59971	Human immune/haema
	26	29.8	29.5	1688	22	AAH34940	Human colon cancer
	27	29.4	29.1	1951	22	ABA44228	Human breast cell
	28	29.4	29.1	1951	22	ABA54678	Human foetal liver
	29	29.4	29.1	1951	22	ABA24460	Probe #2926 for ge
	30	29.4	29.1	1951	22	AAK02965	Human brain expres
	31	29.4	29.1	1951	22	AAK28411	Human bone marrow
	32	29.4	29.1	1951	22	AAI12976	Probe #2909 for ge
	33	29.4	29.1	1951	22	AAI34335	Probe #3021 used t
	34	29.4	29.1	1951	22	AAI02894	Probe #2885 used t
	35	28.8	28.5	17084	22	AAK70764	Human immune/haema
	36	28.8	28.5	17084	22	AAK71755	Human immune/haema
c	37	28.6	28.3	19062	21	AAF21280	Human low adenosin
c	38	28.6	28.3	23142	21	AAA35158	Human adenosine re
c	39	28.6	28.3	30417	21	AAF21282	Human low adenosin
c	40	28.6	28.3	30417	21	AAA35160	Human adenosine re
c	41	28.4	28.1	9674	22	AAL02893	Human reproductive
c	42	28.4	28.1	20387	19	AAV62159	HSV-2 strain SB5 C
c	43	28.4	28.1	26338	19	AAV62134	HSV-2 strain SB5 C
c	44	28.4	28.1	117213	19	AAV62176	HSV-2 strain SB5 C
c	45	28.4	28.1	154746	24	AAD25519	Human herpesvirus

ALIGNMENTS

RESULT 1
AAH43683
ID AAH43683 standard; DNA; 1722 BP.
XX
AC AAH43683;
XX
DT 21-JAN-2002 (first entry)
XX
DE PRKAG3 intron 4 - intron 10.
XX
KW Human; AMP-activated protein kinase gamma 3 subunit; PRKAG3; variant;
KW metabolic disease; diabetes; obesity; substitution; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT intron 1..13
FT /*tag= a
FT /number= "Intron 4"
FT /note= "3' portion of intron 4"
FT exon 14..95
FT /*tag= b
FT /number= "Exon 5"
FT intron 96..552
FT /*tag= c
FT /number= "Intron 5"
FT exon 553..611
FT /*tag= d
FT /number= "Exon 6"
FT intron 612..736
FT /*tag= e
FT /number= "Intron 6"

```
FT   exon            737..782
FT                   /*tag= f
FT                   /number= "Exon 7"
FT   intron          783..986
FT                   /*tag= g
FT                   /number= "Intron 7"
FT   exon            987..1041
FT                   /*tag= h
FT                   /number= "Exon 8"
FT   intron          1042..1242
FT                   /*tag= i
FT                   /number= "Intron 8"
FT   exon            1243..1369
FT                   /*tag= j
FT                   /number= "Exon 9"
FT   intron          1370..1522
FT                   /*tag= k
FT                   /number= "Intron 9"
FT   exon            1523..1688
FT                   /*tag= l
FT                   /number= "Exon 10"
FT   intron          1689..1722
FT                   /*tag= i
FT                   /number= "Intron 10"
FT                   /note= "5' portion of intron 10"
XX
PN   WO200177305-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-SE00765.
XX
PR   07-APR-2000; 2000US-195665P.
XX
PA   (AREX-) AREXIS AB.
XX
PI   Andersson L,  Luthman H,  Marklund S;
XX
DR   WPI; 2001-657170/75.
XX
PT   New variants of human AMP-activated protein kinase gamma3 subunit
PT   associated with a metabolic disease e.g. diabetes or obesity and method
PT   for determining a risk estimate of diseases in subject by detecting the
PT   variant -
XX
PS   Example 1; Fig 3; 25pp; English.
XX
CC   The sequences given in AAH43681-84 represents genomic fragments
CC   encoding the human AMP-activated protein kinase gamma 3 subunit
CC   (PRKAG3). Detecting the presence of the PRKAG3 DNA, or a variant,
CC   is useful in determining a risk estimate of a metabolic disease,
CC   such as diabetes or obesity, in a subject.  The variation may occur
CC   in exons 3, 4 or 10.  In exon 3 variation may be a substitution of
CC   a G for a C at nucleotide 320, resulting in the amino acid
CC   substitution P71A; in exon 4 variation may be a substitution of a
CC   T for a C at nucleotide 550; and in exon 10 variation may be a
CC   substitution of a T for a C at nucleotide 1037, resulting in the
CC   amino acid substitution R340W.  There may also be nucleotide variation
CC   in intron 6.
XX
SQ   Sequence 1722 BP; 321 A; 504 C; 534 G; 363 T; 0 other;


  Query Match             100.0%;  Score 101;  DB 22;  Length 1722;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-18;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 aggtccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcg 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     600 aggtccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcg 659

Qy      61 gaggggagtcctcctggagcctggtgccctagaagcccacg 101
           |||||||||||||||||||||||||||||||||||||||||
Db     660 gaggggagtcctcctggagcctggtgccctagaagcccacg 700


RESULT   2
ABA61253
ID   ABA61253 standard; DNA; 550 BP.
XX
AC   ABA61253;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #9558.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00669.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human fetal liver -
XX
PS   Claim 1; SEQ ID NO 9558; 639pp + sequence listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for
CC   measuring human gene expression in a sample derived from human foetal
CC   liver. The single exon nucleic acid probes may be used for predicting,
CC   measuring and displaying gene expression in samples derived from human
CC   fetal liver. The present sequence is a single exon nucleic acid
CC   probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 550 BP; 154 A; 139 C; 127 G; 130 T; 0 other;


  Query Match             32.3%;  Score 32.6;  DB 22;  Length 550;
  Best Local Similarity   60.9%;  Pred. No. 4.3;
  Matches   53;  Conservative    0;  Mismatches   34;  Indels    0;  Gaps    0;

Qy      10 ctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagt 69
           ||| ||| || | ||| |||       || |   ||||| |||| |||| |  ||||||
Db     185 ctgatgaagactaggcgggggcaaggctgtggcaccagatgggcaggggtgctggggagg 244

Qy      70 cctcctggagcctggtgccctagaagc 96
           ||    | | || | | ||||  || |
Db     245 ccaaaagtaaccagcttccctccaaac 271


RESULT   3
ABA29092
ID   ABA29092 standard; DNA; 550 BP.
XX
AC   ABA29092;
```

```
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #7558 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00666.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts -
XX
PS   Claim 1; SEQ ID No 7558; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
CC   from the human heart via microarrays. By measuring gene expression, the
CC   probes are useful for predicting, diagnosing, grading, staging,
CC   monitoring and prognosing diseases of the human heart and vascular system
CC   e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC   congenital heart disease.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 550 BP; 154 A; 139 C; 127 G; 130 T; 0 other;


  Query Match             32.3%;  Score 32.6;  DB 22;  Length 550;
  Best Local Similarity   60.9%;  Pred. No. 4.3;
  Matches   53;  Conservative    0;  Mismatches   34;  Indels    0;  Gaps    0;

Qy      10 ctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagt 69
           ||| ||| || | ||| |||       || |   ||||| |||| |||| |  ||||||
Db     185 ctgatgaagactaggcgggggcaaggctgtggcaccagatgggcaggggtgctggggagg 244

Qy      70 cctcctggagcctggtgccctagaagc 96
           ||    | | || | | ||||  || |
Db     245 ccaaaagtaaccagcttccctccaaac 271


RESULT   4
AAK09549
ID   AAK09549 standard; DNA; 550 BP.
XX
AC   AAK09549;
XX
DT   05-NOV-2001  (first entry)
XX
DE   Human brain expressed single exon probe SEQ ID NO: 9540.
XX
KW   Human; brain expressed exon; gene expression analysis; probe;
KW   microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW   epilepsy; cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157275-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00667.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483446/52.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   brains -
XX
PS   Example 4; SEQ ID NO: 9540; 650pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   brain. They can be used to measure gene expression in brain cell samples,
CC   which may enable the diagnosis and improved treatment of nervous system
CC   diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC   epilepsy and cancers. The present sequence is one of the probes of the
CC   invention.
XX
SQ   Sequence 550 BP; 154 A; 139 C; 127 G; 130 T; 0 other;


  Query Match             32.3%;  Score 32.6;  DB 22;  Length 550;
  Best Local Similarity   60.9%;  Pred. No. 4.3;
  Matches   53;  Conservative    0;  Mismatches   34;  Indels    0;  Gaps    0;

Qy      10 ctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagt 69
           ||| ||| || | ||| |||       || |   ||||| |||| |||| |  ||||||
Db     185 ctgatgaagactaggcgggggcaaggctgtggcaccagatgggcaggggtgctggggagg 244

Qy      70 cctcctggagcctggtgccctagaagc 96
           ||    | | || | | ||||  || |
Db     245 ccaaaagtaaccagcttccctccaaac 271


RESULT   5
AAK35441
ID   AAK35441 standard; DNA; 550 BP.
XX
AC   AAK35441;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human bone marrow expressed single exon probe SEQ ID NO: 9998.
XX
KW   Human; bone marrow expressed exon; gene expression analysis; probe;
KW   microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157276-A2.
```

```
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00668.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488900/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human bone marrow -
XX
PS   Example 4; SEQ ID NO: 9998; 658pp + Sequence Listing; English.
XX
CC   The present invention provides a number of single exon nucleic acid
CC   probes which are derived from genomic sequences expressed in the human
CC   bone marrow. They can be used to measure gene expression in bone marrow
CC   samples, which may enable the improved diagnosis and treatment of cancers
CC   such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC   the probes of the invention.
XX
SQ   Sequence 550 BP; 154 A; 139 C; 127 G; 130 T; 0 other;


  Query Match             32.3%;  Score 32.6;  DB 22;  Length 550;
  Best Local Similarity   60.9%;  Pred. No. 4.3;
  Matches   53;  Conservative    0;  Mismatches   34;  Indels    0;  Gaps    0;

Qy      10 ctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagt 69
           ||| ||| || | ||| |||       || |   ||||| |||| |||| |  ||||||
Db     185 ctgatgaagactaggcgggggcaaggctgtggcaccagatgggcaggggtgctggggagg 244

Qy      70 cctcctggagcctggtgccctagaagc 96
           ||    | | || | | ||||  || |
Db     245 ccaaaagtaaccagcttccctccaaac 271


RESULT   6
AAI17045
ID   AAI17045 standard; DNA; 550 BP.
XX
AC   AAI17045;
XX
DT   12-OCT-2001  (first entry)
XX
DE   Probe #6978 for gene expression analysis in human cervical cell sample.
XX
KW   Probe; human; microarray; gene expression; cervical epithelial cell;
KW   cervical cancer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488901/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human cervical epithelial cells -
XX
PS   Claim 25; SEQ ID No 6978; 487pp; English.
XX
CC   The present invention relates to human single exon nucleic acid probes
CC   (SENP). The present sequence is one such probe. The SENPs are derived
CC   from human HeLa cells. The SENPs can be used to produce a single exon
CC   microarray, which can be used for measuring human gene expression in a
CC   sample derived from human cervical epithelial cells. By measuring gene
CC   expression, the probes are therefore useful in grading and/or staging
CC   of diseases of the cervix, notably cervical cancer.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 550 BP; 154 A; 139 C; 127 G; 130 T; 0 other;


  Query Match             32.3%;  Score 32.6;  DB 22;  Length 550;
  Best Local Similarity   60.9%;  Pred. No. 4.3;
  Matches   53;  Conservative    0;  Mismatches   34;  Indels    0;  Gaps    0;

Qy      10 ctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagt 69
           ||| ||| || | ||| |||       || |   ||||| |||| |||| |  ||||||
Db     185 ctgatgaagactaggcgggggcaaggctgtggcaccagatgggcaggggtgctggggagg 244

Qy      70 cctcctggagcctggtgccctagaagc 96
           ||    | | || | | ||||  || |
Db     245 ccaaaagtaaccagcttccctccaaac 271


RESULT   7
AAI41154
ID   AAI41154 standard; DNA; 550 BP.
XX
AC   AAI41154;
XX
DT   17-OCT-2001  (first entry)
XX
DE   Probe #9840 used to measure gene expression in human placenta sample.
XX
KW   Probe; microarray; human; placenta; antenatal diagnosis;
KW   genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157272-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00663.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
```

```
XX
DR   WPI; 2001-488897/53.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human placenta -
XX
PS   Claim 25; SEQ ID No 9840; 654pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes (SENP).
CC   The present sequence is one such probe. The probes are useful for
CC   producing a microarray for predicting, measuring and displaying gene
CC   expression in samples derived from human placenta. The probes are useful
CC   for antenatal diagnosis of human genetic disorders.
XX
SQ   Sequence 550 BP; 154 A; 139 C; 127 G; 130 T; 0 other;


  Query Match             32.3%;  Score 32.6;  DB 22;  Length 550;
  Best Local Similarity   60.9%;  Pred. No. 4.3;
  Matches   53;  Conservative    0;  Mismatches   34;  Indels    0;  Gaps    0;

Qy      10 ctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagt 69
           ||| ||| || | ||| |||       || |   ||||| |||| |||| |  ||||||
Db     185 ctgatgaagactaggcgggggcaaggctgtggcaccagatgggcaggggtgctggggagg 244

Qy      70 cctcctggagcctggtgccctagaagc 96
           ||    | | || | | ||||  || |
Db     245 ccaaaagtaaccagcttccctccaaac 271


RESULT   8
AAK76637
ID   AAK76637 standard; DNA; 680 BP.
XX
AC   AAK76637;
XX
DT   07-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:31449.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01354.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
```

```
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
PR   08-NOV-2000; 2000US-0246528.
PR   08-NOV-2000; 2000US-0246532.
'R   08-NOV-2000; 2000US-0246609.
PR   08-NOV-2000; 2000US-0246610.
PR   08-NOV-2000; 2000US-0246611.
PR   08-NOV-2000; 2000US-0246613.
PR   17-NOV-2000; 2000US-0249207.
PR   17-NOV-2000; 2000US-0249208.
PR   17-NOV-2000; 2000US-0249209.
PR   17-NOV-2000; 2000US-0249210.
PR   17-NOV-2000; 2000US-0249211.
PR   17-NOV-2000; 2000US-0249212.
PR   17-NOV-2000; 2000US-0249213.
PR   17-NOV-2000; 2000US-0249214.
PR   17-NOV-2000; 2000US-0249215.
PR   17-NOV-2000; 2000US-0249216.
PR   17-NOV-2000; 2000US-0249217.
PR   17-NOV-2000; 2000US-0249218.
PR   17-NOV-2000; 2000US-0249244.
PR   17-NOV-2000; 2000US-0249245.
PR   17-NOV-2000; 2000US-0249264.
PR   17-NOV-2000; 2000US-0249265.
PR   17-NOV-2000; 2000US-0249297.
PR   17-NOV-2000; 2000US-0249299.
PR   17-NOV-2000; 2000US-0249300.
PR   01-DEC-2000; 2000US-0250160.
PR   01-DEC-2000; 2000US-0250391.
PR   05-DEC-2000; 2000US-0251030.
PR   05-DEC-2000; 2000US-0251988.
PR   05-DEC-2000; 2000US-0256719.
PR   06-DEC-2000; 2000US-0251479.
PR   08-DEC-2000; 2000US-0251856.
PR   08-DEC-2000; 2000US-0251868.
PR   08-DEC-2000; 2000US-0251869.
PR   08-DEC-2000; 2000US-0251989.
PR   08-DEC-2000; 2000US-0251990.
PR   11-DEC-2000; 2000US-0254097.
PR   05-JAN-2001; 2001US-0259678.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Barash SC,  Ruben SM;
XX
DR   WPI; 2001-483426/52.
XX
PT   Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT   useful for preventing, diagnosing and/or treating cancers and
PT   metastasis -
XX
PS   Disclosure; SEQ ID NO 31449; 3071pp + Sequence Listing; English.
XX
CC   AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC   amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC   activity, and can be used in gene therapy and vaccine production. (I)
CC   proteins and polynucleotides may be used in the prevention, diagnosis and
CC   treatment of diseases associated with inappropriate (I) expression. For
CC   example, they may be used to treat disorders associated with decreased
CC   expression by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of (I) by expressing inactive proteins or to
CC   supplement the patients own production of (I). Additionally, (I)
CC   polynucleotides may be used to produce the secreted (I), by inserting
CC   the nucleic acids into a host cell and culturing the cell to express the
CC   protein. (I) proteins and polynucleotides may be used to prevent,
CC   diagnose and treat immune/haematopoietic-related diseases, especially
CC   cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC   to AAK87694 represent human immune/haematopoietic antigen genomic
CC   sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC   represent sequences used in the exemplification of the present invention.
XX
SQ   Sequence 680 BP; 116 A; 213 C; 185 G; 166 T; 0 other;
```

```
  Query Match             32.3%;  Score 32.6;  DB 22;  Length 680;
  Best Local Similarity   58.9%;  Pred. No. 4.3;
  Matches   56;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy       7 cccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggagggg 66
           ||||||  | ||||||||| ||||  |||   | ||  || |   ||| ||  |   |
Db     569 cccctgcaggggagtggggtgggtctctcgagtgcggcctgccaggcaggaacttggctc 628

Qy      67 agtcctcctggagcctggtgccctagaagcccacg 101
           |   | | || ||||  |   ||| | |||||| |
Db     629 atagcccatgcagccatgcagccttgcagcccagg 663


RESULT   9
AAS81565/c
ID   AAS81565 standard; cDNA; 3544 BP.
XX
AC   AAS81565;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #17369.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG17378.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 1; SEQ ID No 17369; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
```

```
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 3544 BP; 848 A; 1024 C; 1040 G; 632 T; 0 other;


  Query Match           32.3%;  Score 32.6;  DB 23;  Length 3544;
  Best Local Similarity  60.9%;  Pred. No. 4.7;
  Matches   53;  Conservative    0;  Mismatches   34;  Indels    0;  Gaps    0;

Qy       4 tccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggagg 63
           ||  | |||| | | || ||||||  |  |||  ||||  |  |    ||  || |  |
Db    2918 TCTGCACTGGGGCGTAGGGGGCTGCCAGGCTTGGGGGCCGCAGGGTCAGCTTGGCCTCAC 2859

Qy      64 gggagtcctcctggagcctggtgccct 90
           ||   |||||||||| ||||  || ||
Db    2858 GGTGCTCCTCCTGGATCCTGAAGCGCT 2832


RESULT  10
AAS74057/c
ID   AAS74057 standard; cDNA; 3750 BP.
XX
AC   AAS74057;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #9861.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG09870.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 1; SEQ ID No 9861; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 3750 BP; 841 A; 1043 C; 966 G; 900 T; 0 other;


  Query Match           32.1%;  Score 32.4;  DB 23;  Length 3750;
  Best Local Similarity  60.0%;  Pred. No. 5.4;
  Matches   54;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy      10 ctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggaagt 69
           ||| ||| || | ||| |||       || |   ||||| |||| |||| |  ||||||
Db     475 CTGATGAAGACTAGGCGGGGGCAAGGCTGTGGCACCAGATGGGCAGGGGTGCTGGGGAGG 416

Qy      70 cctcctggagcctggtgccctagaagccca 99
           ||    | | || | | ||||  ||   ||
Db     415 CCAAAAGTAACCAGCTTCCCTCCAAAAGCA 386


RESULT  11
AAH90107/c
ID   AAH90107 standard; cDNA; 3889 BP.
XX
AC   AAH90107;
XX
DT   01-OCT-2001  (first entry)
XX
DE   Human bone marrow cDNA, SEQ ID NO: 464.
XX
KW   Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
KW   antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
KW   immunosuppressive; gene therapy; cytokine cell proliferation;
KW   cell differentiation modulator; immune disorder; infection; cancer;
KW   human immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153453-A2.
XX
PD   26-JUL-2001.
XX
PF   23-DEC-2000; 2000WO-US34960.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   30-NOV-2000; 2000US-0250583.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Ford JE,  Boyle BJ,  Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y;
PI   Ren F,  Wang J,  Werhman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Drmanac RT;
XX
DR   WPI; 2001-488707/53.
DR   P-PSDB; AAM00988.
XX
PT   Novel bone-marrow-expressed polynucleotides and polypeptides, useful
PT   for treating e.g. cancer and immune deficiency disorders -
XX
PS   Claim 1; Page 574-579; 648pp; English.
XX
CC   The present sequence is one of 251 novel human polynucleotides
CC   expressed in the bone marrow. The polynucleotide and the
CC   polypeptide encoded by it are useful in the treatment of various
CC   immune deficiencies and disorders. The deficiencies and disorders may
CC   be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
```

```
CC   infection, or may result from an autoimmune disorder, a coagulation
CC   disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
CC   suppression of an inflammatory response or treatment of a nervous
CC   system disorder such as Alzheimer's disease. Detection of the presence
CC   or increased expression of the polynucleotide or the protein it
CC   encodes is useful for the diagnosis and/or prognosis of one
CC   or more types of cancer. The polynucleotide and polypeptide can be
CC   used as nutritional sources or supplements and in the screening of
CC   chemical compounds as potential drugs.
XX
SQ   Sequence 3889 BP; 861 A; 1090 C; 1008 G; 930 T; 0 other;


  Query Match             32.1%;  Score 32.4;  DB 22;  Length 3889;
  Best Local Similarity   60.0%;  Pred. No. 5.4;
  Matches   54;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy      10 ctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagt 69
           ||| ||| || | ||| |||       || |   ||||| |||| |||| |  ||||||
Db     633 CTGATGAAGACTAGGCGGGGGCAAGGCTGTGGCACCAGATGGGCAGGGGTGCTGGGGAGG 574

Qy      70 cctcctggagcctggtgccctagaagccca 99
           ||    | | || | | ||||  ||   ||
Db     573 CCAAAAGTAACCAGCTTCCCTCCAAAAGCA 544


RESULT  12
AAH90056/c
ID   AAH90056 standard; cDNA; 3892 BP.
XX
AC   AAH90056;
XX
DT   01-OCT-2001  (first entry)
XX
DE   Human bone marrow cDNA, SEQ ID NO: 300.
XX
KW   Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
KW   antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
KW   immunosuppressive; gene therapy; cytokine cell proliferation;
KW   cell differentiation modulator; immune disorder; infection; cancer;
KW   human immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153453-A2.
XX
PD   26-JUL-2001.
XX
PF   23-DEC-2000; 2000WO-US34960.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   30-NOV-2000; 2000US-0250583.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Ford JE,  Boyle BJ,  Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y;
PI   Ren F,  Wang J,  Werhman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Drmanac RT;
XX
DR   WPI; 2001-488707/53.
DR   P-PSDB; AAM00937.
XX
PT   Novel bone-marrow-expressed polynucleotides and polypeptides, useful
PT   for treating e.g. cancer and immune deficiency disorders -
XX
PS   Claim 1; Page 411-412; 648pp; English.
XX
CC   The present sequence is one of 251 novel human polynucleotides
CC   expressed in the bone marrow. The polynucleotide and the
CC   polypeptide encoded by it are useful in the treatment of various
CC   immune deficiencies and disorders. The deficiencies and disorders may
CC   be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
CC   infection, or may result from an autoimmune disorder, a coagulation
CC   disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
CC   suppression of an inflammatory response or treatment of a nervous
CC   system disorder such as Alzheimer's disease. Detection of the presence
CC   or increased expression of the polynucleotide or the protein it
CC   encodes is useful for the diagnosis and/or prognosis of one
CC   or more types of cancer. The polynucleotide and polypeptide can be
CC   used as nutritional sources or supplements and in the screening of
CC   chemical compounds as potential drugs.
XX
SQ   Sequence 3892 BP; 860 A; 1091 C; 1009 G; 932 T; 0 other;


  Query Match             32.1%;  Score 32.4;  DB 22;  Length 3892;
  Best Local Similarity   60.0%;  Pred. No. 5.4;
  Matches   54;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy      10 ctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagt 69
           ||| ||| || | ||| |||       || |   ||||| |||| |||| |  ||||||
Db     633 CTGATGAAGACTAGGCGGGGGCAAGGCTGTGGCACCAGATGGGCAGGGGTGCTGGGGAGG 574

Qy      70 cctcctggagcctggtgccctagaagccca 99
           ||    | | || | | ||||  ||   ||
Db     573 CCAAAAGTAACCAGCTTCCCTCCAAAAGCA 544


RESULT  13
AAH89943/c
ID   AAH89943 standard; cDNA; 4886 BP.
XX
AC   AAH89943;
XX
DT   01-OCT-2001  (first entry)
XX
DE   Human bone marrow cDNA, SEQ ID NO: 74.
XX
KW   Human; bone marrow; antiinflammatory; cytostatic; neuroprotective;
KW   antiviral; antibacterial; antifungal; anti-HIV; haemostatic;
KW   immunosuppressive; gene therapy; cytokine cell proliferation;
KW   cell differentiation modulator; immune disorder; infection; cancer;
KW   human immunodeficiency virus; HIV; autoimmune disorder; haemophilia; ss.
XX
OS   Homo sapiens.
XX
PN   WO200153453-A2.
XX
PD   26-JUL-2001.
XX
PF   23-DEC-2000; 2000WO-US34960.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   30-NOV-2000; 2000US-0250583.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Ford JE,  Boyle BJ,  Tang YT,  Liu C,  Asundi V,  Chen R,  Ma Y;
PI   Ren F,  Wang J,  Werhman T,  Xu C,  Xue AJ,  Yang Y,  Zhang J;
PI   Zhao QA,  Zhou P,  Drmanac RT;
XX
DR   WPI; 2001-488707/53.
DR   P-PSDB; AAM00824.
XX
```

```
PT   Novel bone-marrow-expressed polynucleotides and polypeptides, useful
PT   for treating e.g. cancer and immune deficiency disorders -
XX
PS   Claim 1; Page 259-260; 648pp; English.
XX
CC   The present sequence is one of 251 novel human polynucleotides
CC   expressed in the bone marrow. The polynucleotide and the
CC   polypeptide encoded by it are useful in the treatment of various
CC   immune deficiencies and disorders. The deficiencies and disorders may
CC   be genetic, may be caused by a viral (e.g. HIV), bacterial or fungal
CC   infection, or may result from an autoimmune disorder, a coagulation
CC   disorder (e.g. haemophilia), inhibition of tumour cell proliferation,
CC   suppression of an inflammatory response or treatment of a nervous
CC   system disorder such as Alzheimer's disease. Detection of the presence
CC   or increased expression of the polynucleotide or the protein it
CC   encodes is useful for the diagnosis and/or prognosis of one
CC   or more types of cancer. The polynucleotide and polypeptide can be
CC   used as nutritional sources or supplements and in the screening of
CC   chemical compounds as potential drugs.
XX
SQ   Sequence 4886 BP; 1117 A; 1370 C; 1266 G; 1133 T; 0 other;


  Query Match             32.1%;  Score 32.4;  DB 22;  Length 4886;
  Best Local Similarity   60.0%;  Pred. No. 5.5;
  Matches   54;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy      10 ctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggagggagt 69
           ||| ||| || | ||| |||       || |   ||||| |||| |||| |  ||||||
Db    1616 CTGATGAAGACTAGGCGGGGGCAAGGCTGTGGCACCAGATGGGCAGGGGTGCTGGGGAGG 1557

Qy      70 cctcctggagcctggtgccctagaagccca 99
           ||    | | || | | |||| ||   ||
Db    1556 CCAAAAGTAACCAGCTTCCCTCCAAAAGCA 1527


RESULT 14
ABA73753
ID   ABA73753 standard; DNA; 231 BP.
XX
AC   ABA73753;
XX
DT   01-FEB-2002  (first entry)
XX
DE   Human foetal liver single exon nucleic acid probe #22058.
XX
KW   Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157277-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00669.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-483447/52.
XX
PT   Human genome-derived single exon nucleic acid probes useful for
PT   analyzing gene expression in human fetal liver -
XX
PS   Claim 4; SEQ ID NO 22058; 639pp + sequence listing; English.
XX
CC   The invention relates to a single exon nucleic acid probe for
CC   measuring human gene expression in a sample derived from human foetal
CC   liver. The single exon nucleic acid probes may be used for predicting,
CC   measuring and displaying gene expression in samples derived from human
CC   fetal liver. The present sequence is a single exon nucleic acid
CC   probe of the invention.
CC   Note: The sequence data for this patent did not form part of the
CC   printed specification, but was obtained in electronic format directly
CC   from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 231 BP; 54 A; 57 C; 63 G; 57 T; 0 other;


  Query Match             31.9%;  Score 32.2;  DB 22;  Length 231;
  Best Local Similarity   61.2%;  Pred. No. 5.3;
  Matches   52;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy      10 ctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggagggagt 69
           ||| ||| || | ||| |||       || |   ||||| |||| |||| |  ||||||
Db    146 ctgatgaagactaggcgggggcaaggctgtggcaccagatgggcaggggtgctggggagg 205

Qy      70 cctcctggagcctggtgccctagaa 94
           ||    | | || | | ||||  ||
Db    206 ccaaaagtaaccagcttccctccaa 230


RESULT 15
ABA38943
ID   ABA38943 standard; DNA; 231 BP.
XX
AC   ABA38943;
XX
DT   23-JAN-2002  (first entry)
XX
DE   Probe #17409 for gene expression analysis in human heart cell sample.
XX
KW   Human; gene expression; heart; microarray; vascular system; probe;
KW   cardiovascular disease; hypertension; cardiac arrhythmia;
KW   congenital heart disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157274-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00666.
XX
PR   04-FEB-2000; 2000US-0180312.
PR   26-MAY-2000; 2000US-0207456.
PR   30-JUN-2000; 2000US-0608408.
PR   03-AUG-2000; 2000US-0632366.
PR   21-SEP-2000; 2000US-0234687.
PR   27-SEP-2000; 2000US-0236359.
PR   04-OCT-2000; 2000GB-0024263.
XX
PA   (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI   Penn SG,  Hanzel DK,  Chen W,  Rank DR;
XX
DR   WPI; 2001-488899/53.
XX
PT   Single exon nucleic acid probes for analyzing gene expression in human
PT   hearts -
XX
PS   Claim 4; SEQ ID No 17409; 530pp; English.
XX
CC   The present invention relates to single exon nucleic acid probes for
CC   measuring human gene expression in a sample derived from human heart. The
CC   present sequence is one such probe. The probes may be used for
CC   predicting, measuring and displaying gene expression in samples derived
```

CC from the human heart via microarrays. By measuring gene expression, the
CC probes are useful for predicting, diagnosing, grading, staging,
CC monitoring and prognosing diseases of the human heart and vascular system
CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
CC congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 231 BP; 54 A; 57 C; 63 G; 57 T; 0 other;

Query Match 31.9%; Score 32.2; DB 22; Length 231;
Best Local Similarity 61.2%; Pred. No. 5.3;
Matches 52; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

```
Qy      10 ctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagt 69
           ||| ||| || | ||| |||       || |   ||||| |||| |||| |  ||||||
Db     146 ctgatgaagactaggcgggggcaaggctgtggcaccagatgggcaggggtgctggggagg 205

Qy      70 cctcctggagcctggtgccctagaa 94
           ||    | | || | | ||||  ||
Db     206 ccaaaagtaaccagcttccctccaa 230
```

Search completed: October 3, 2002, 16:30:18
Job time: 14288 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 16:14:46 ; Search time 5701.1 Seconds
(without alignments)
239.110 Million cell updates/sec

Title: US-09-826-581-3_COPY_600_700
Perfect score: 101
Sequence: 1 aggtccccccctggtgaggag.....tggtgccctagaagcccacg 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	36	35.6	682	10	BI960023	BI960023 HVSMEn002
	2	35.2	34.9	833	10	BE559614	BE559614 601347308
	3	35	34.7	377	9	AW315288	AW315288 12477 MAR
	4	33	32.7	872	9	AL530561	AL530561 AL530561
	5	32.8	32.5	656	10	BI960536	BI960536 HVSMEn002
	6	32.6	32.3	282	9	AA872225	AA872225 oh71b02.s
c	7	32.6	32.3	354	9	AW026388	AW026388 wv12f05.x
c	8	32.6	32.3	539	10	BF808916	BF808916 PM4-CI015
	9	32.2	31.9	685	9	AL585199	AL585199 AL585199
	10	32	31.7	735	10	BG309093	BG309093 HVSMEc000
c	11	31.8	31.5	382	10	BE600720	BE600720 PI1_91_G1
c	12	31.8	31.5	463	10	BE600139	BE600139 PI1_79_G0
c	13	31.8	31.5	474	10	BE598197	BE598197 PI1_66_H0
c	14	31.8	31.5	487	10	BE600794	BE600794 PI1_90_C0
c	15	31.8	31.5	532	10	BE595429	BE595429 PI1_55_D0
	16	31.6	31.3	695	12	AG083214	AG083214 Pan trogl
c	17	31.6	31.3	889	10	BI951430	BI951430 HVSME1002
c	18	31.4	31.1	452	9	AI863974	AI863974 wj54c05.x
c	19	31.2	30.9	407	9	AU162756	AU162756 AU162756
c	20	31.2	30.9	516	9	AU162755	AU162755 AU162755
	21	31.2	30.9	539	9	AL514039	AL514039 AL514039
	22	31.2	30.9	688	9	AL514251	AL514251 AL514251
	23	31.2	30.9	883	12	AZ527770	AZ527770 ENTDA48TF
c	24	31	30.7	242	10	BF685935	BF685935 602140240
	25	31	30.7	469	9	AI087063	AI087063 oy70d09.x
c	26	31	30.7	526	10	BE906990	BE906990 601499803
c	27	31	30.7	529	10	BG481234	BG481234 602528560
c	28	31	30.7	531	9	BE019039	BE019039 bb46b11.y
c	29	31	30.7	535	10	BE250790	BE250790 601106982
c	30	31	30.7	535	10	BE296883	BE296883 601176963
c	31	31	30.7	537	10	BE294571	BE294571 601176144
c	32	31	30.7	538	10	BE298630	BE298630 601119416
c	33	31	30.7	542	10	BE295054	BE295054 601173761
c	34	31	30.7	553	10	BI160864	BI160864 602864791
c	35	31	30.7	557	10	BI767103	BI767103 603054158
c	36	31	30.7	572	10	BE783728	BE783728 601471127
c	37	31	30.7	581	9	AL538073	AL538073 AL538073
c	38	31	30.7	606	9	AW963549	AW963549 EST375622
c	39	31	30.7	610	10	BE300397	BE300397 600944511
c	40	31	30.7	611	10	BI254382	BI254382 602976691
c	41	31	30.7	617	10	BE271259	BE271259 600943675
c	42	31	30.7	629	10	BF794131	BF794131 602255503
c	43	31	30.7	640	9	AL531142	AL531142 AL531142
c	44	31	30.7	649	9	AL531376	AL531376 AL531376
c	45	31	30.7	658	10	BF982575	BF982575 602305702

ALIGNMENTS

RESULT 1
BI960023
LOCUS BI960023 682 bp mRNA linear EST 22-OCT-2001
DEFINITION HVSMEn0022O08f Hordeum vulgare rachis EST library HVcDNA0015 (normal) Hordeum vulgare cDNA clone HVSMEn0022O08f, mRNA sequence.
ACCESSION BI960023
VERSION BI960023.1 GI:16311278
KEYWORDS EST.
SOURCE barley.
ORGANISM Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae ; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 682)
AUTHORS Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Chin,A., Begum,D., Frisch,D., Atkins,M., Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons ,J., Oates,R. and Main,D.
TITLE Development of a genetically and physically anchored EST resource for barley genomics: Morex rachis cDNA library
JOURNAL Unpublished (2001)
COMMENT Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hq bases = 121
Seq primer: AATTAACCCTCACTAAAGGG
High quality sequence start: 2
High quality sequence stop: 553.
FEATURES Location/Qualifiers
source 1. .682
/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSMEn0022O08f"
/clone_lib="Hordeum vulgare rachis EST library HVcDNA0015 (normal)"
/tissue_type="Rachis"

```
                     /lab_host="TJC121"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; Plants were grown at Washington State University,
                     Pullman, WA in a greenhouse, the rachises were excised and
                     frozen in liquid nitrogen (Kleinhofs lab). In the TJ Close
                     lab at the University of California, Riverside total RNA
                     was prepared, poly(A) was purified, one primary
                     unamplified cDNA library was made, and 1 million pfu were
                     in vivo excised to give pBluescript SK(-) cDNA phagemids
                     (Chin). Phagemids were plated and picked at the Clemson
                     University Genomics Institute (CUGI) (Begum, Palmer,
                     Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
                     sequencing and sequence analysis were performed at CUGI
                     (Wing, Yu, Frisch, Henry, Simmons, Rambo, Main). The
                     sequence has been trimmed to remove vector sequence and
                     contains a minimum of 100 bases of phred value 20 or
                     above. For more details on library preparation and
                     sequence analysis see
                     http://www.genome.clemson.edu/projects/barley. To order
                     this clone see http://www.genome.clemson.edu/orders Also
                     see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
                     Genetically and physically anchored EST resources for
                     barley genomics. Barley Genetics Newsletter 31:29-30.
                     (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"
BASE COUNT       75 a    262 c    174 g    171 t
ORIGIN


  Query Match           35.6%;  Score 36;  DB 10;  Length 682;
  Best Local Similarity  60.0%;  Pred. No. 47;
  Matches   60;  Conservative    0;  Mismatches   40;  Indels    0;  Gaps    0;

Qy       1 aggtccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcg 60
           ||| | || | |||| |||| |||| | ||| || ||  ||  | |||  |||| || ||
Db      17 AGGGCGCCGCCGGTGGGGAGGGGCGGCGAAGCTGATCTGCGACGAGAATGGCGTGGACG 76

Qy      61 gaggggagtcctcctggagcctggtgccctagaagcccac 100
             |||    ||| || | | | |  |    |   ||| | |
Db      77 AGGGGTTCCCCTACTCGCGGCCGTCGGTGTGCGAGCTCTC 116


RESULT   2
BE559614
LOCUS       BE559614                 833 bp    mRNA    linear   EST 15-AUG-2000
DEFINITION  601347308F1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:3688176 5',
            mRNA sequence.
ACCESSION   BE559614
VERSION     BE559614.1  GI:9803323
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 833)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM380  row: p  column: 01
            High quality sequence stop: 687.
FEATURES             Location/Qualifiers
     source          1. .833
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3688176"
                     /clone_lib="NIH_MGC_8"
                     /tissue_type="Burkitt lymphoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      215 a    175 c    257 g    186 t
ORIGIN


  Query Match           34.9%;  Score 35.2;  DB 10;  Length 833;
  Best Local Similarity  65.0%;  Pred. No. 74;
  Matches   52;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy       7 cccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggagggg 66
           |||||||   ||| |||| |||| ||  ||| | |||  | ||||||| || |  ||||
Db     739 CCCCTGGGTTGGATTGGGGTGGGGGTCGAATGCGGACCTCGGGGGGCGGCGGGGCTGGGG 798

Qy      67 agtcctcctggagcctggtg 86
           |   | |  ||   || |||
Db     799 AAGGCCCTGGGGCGCTTGTG 818


RESULT   3
AW315288
LOCUS       AW315288                 377 bp    mRNA    linear   EST 25-APR-2001
DEFINITION  12477 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   AW315288
VERSION     AW315288.1  GI:6744544
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 377)
  AUTHORS   Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
            ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
            Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
            Keele,J.W.
  TITLE     Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
  JOURNAL   Genome Res. 11 (4), 626-630 (2001)
  MEDLINE   21180013
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 20
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 142  row: D  column: 8
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .377
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="MARC 2BOV"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled tissue from testis, thymus,
                     semitendonosus muscle, longissimus muscle, pancreas,
```

```
                    adrenal, and endometrium."
BASE COUNT      100 a    119 c     99 g     59 t
ORIGIN


  Query Match            34.7%;  Score 35;  DB 9;  Length 377;
  Best Local Similarity  59.6%;  Pred. No. 76;
  Matches   59;  Conservative    0;  Mismatches   40;  Indels    0;  Gaps    0;

Qy       2 ggtccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcgg 61
           || |||  ||  | || ||  |  | | || |  | | || ||||    |||||| | |
Db     276 GGACCCTTCTTCTTAGAAGCAGAATTGCAACCGGAGGAGCTCCCACCCTGGCGGGCGGGC 335

Qy      62 aggggagtcctcctggagcctggtgccctagaagcccac 100
           ||| || ||||| ||||||  ||        |  ||||||
Db     336 GGGGAAGCCCTCCAGGAGCCGGGGAAGGAGGCGGCCCAC 374


RESULT   4
AL530561
LOCUS       AL530561                872 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL530561 LTI_NFL001_NBC4 Homo sapiens cDNA clone CS0DD007YO22 5
            prime, mRNA sequence.
ACCESSION   AL530561
VERSION     AL530561.1  GI:12794054
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 872)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .872
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DD007YO22"
                     /clone_lib="LTI_NFL001_NBC4"
                     /sex="male"
                     /tissue_type="neuroblastoma cells"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector.  Library was normalized. Library was constructed
                     by Life Technologies.  Contact : Feng Liang Life
                     Technologies, a division of Invitrogen 9800 Medical Center
                     Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
                     8371 Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      195 a    189 c    337 g     38 t    113 others
ORIGIN


  Query Match            32.7%;  Score 33;  DB 9;  Length 872;
  Best Local Similarity  52.0%;  Pred. No. 2.4e+02;
  Matches   39;  Conservative   13;  Mismatches   23;  Indels    0;  Gaps    0;

Qy      12 ggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagtcc 71
           || |||||:||    :|||: :  :| :  || ||||:|||:|||| | |:||| : |
Db     378 GGGCAGGAGKGGAAGAKGAAKKKGCKGCKGCCCAAGAGKGGCKGGGGKGCAKGGGCKCCA 437

Qy      72 tcctggagcctggtg 86
           | :|| |   :|:|
Db     438 ACGKGGKGGACKGKG 452
```

```
RESULT   5
BI960536
LOCUS       BI960536                656 bp    mRNA    linear   EST 22-OCT-2001
DEFINITION  HVSMEn0024P07f Hordeum vulgare rachis EST library HVcDNA0015
            (normal) Hordeum vulgare cDNA clone HVSMEn0024P07f, mRNA sequence.
ACCESSION   BI960536
VERSION     BI960536.1  GI:16311791
KEYWORDS    EST.
SOURCE      barley.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 656)
  AUTHORS   Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Chin,A., Begum,D.,
            Frisch,D., Atkins,M., Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons
            ,J., Oates,R. and Main,D.
  TITLE     Development of a genetically and physically anchored EST resource
            for barley genomics: Morex rachis cDNA library
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Total hq bases = 447
            Seq primer: AATTAACCCTCACTAAAGGG
            High quality sequence start: 4
            High quality sequence stop: 541.
FEATURES             Location/Qualifiers
     source          1. .656
                     /organism="Hordeum vulgare"
                     /cultivar="Morex"
                     /db_xref="taxon:4513"
                     /clone="HVSMEn0024P07f"
                     /clone_lib="Hordeum vulgare rachis EST library HVcDNA0015
                     (normal)"
                     /tissue_type="Rachis"
                     /lab_host="TJC121"
                     /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                     XhoI; Plants were grown at Washington State University,
                     Pullman, WA in a greenhouse, the rachises were excised and
                     frozen in liquid nitrogen (Kleinhofs lab). In the TJ Close
                     lab at the University of California, Riverside total RNA
                     was prepared, poly(A) was purified, one primary
                     unamplified cDNA library was made, and 1 million pfu were
                     in vivo excised to give pBluescript SK(-) cDNA phagemids
                     (Chin). Phagemids were plated and picked at the Clemson
                     University Genomics Institute (CUGI) (Begum, Palmer,
                     Frisch, Atkins and Wing). Plasmid DNA preparations, DNA
                     sequencing and sequence analysis were performed at CUGI
                     (Wing, Yu, Frisch, Henry, Simmons, Rambo, Main). The
                     sequence has been trimmed to remove vector sequence and
                     contains a minimum of 100 bases of phred value 20 or
                     above. For more details on library preparation and
                     sequence analysis see
                     http://www.genome.clemson.edu/projects/barley. To order
                     this clone see http://www.genome.clemson.edu/orders Also
                     see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
                     Genetically and physically anchored EST resources for
                     barley genomics. Barley Genetics Newsletter 31:29-30.
                     (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"
BASE COUNT      114 a    215 c    229 g     98 t
ORIGIN


  Query Match            32.5%;  Score 32.8;  DB 10;  Length 656;
  Best Local Similarity  58.0%;  Pred. No. 2.6e+02;
  Matches   58;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;
```

```
Qy       1 aggtccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcg 60
           ||| | || | |||| |||| |||| | ||| || |   ||  | |||  |||| || ||
Db      17 AGGGCGCCGCCGGTGGGGAGGGGCGGCGAAGCTGAGCTGCGACGAGAATGGCGTGGACG 76

Qy      61 gaggggagtcctcctggagcctggtgccctagaagcccac 100
             |||    ||| |  | | | |  |    |   ||| | |
Db      77 AGGGGTTCCCCTACGCGCGGCCGTCGGTGTGCGAGCTCTC 116


RESULT   6
AA872225
LOCUS       AA872225                282 bp    mRNA    linear   EST 29-APR-1998
DEFINITION  oh71b02.s1 NCI_CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1472427 3',
            mRNA sequence.
ACCESSION   AA872225
VERSION     AA872225.1  GI:2968403
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 282)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 874   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 230.
FEATURES             Location/Qualifiers
     source          1. .282
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1472427"
                     /clone_lib="NCI_CGAP_Kid5"
                     /tissue_type="2 pooled tumors (clear cell type)"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     AACTGGAAGAATTCGCGGCCGCAATATTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M. Fatima Bonaldo.  "
BASE COUNT       49 a    107 c     71 g     55 t
ORIGIN


  Query Match             32.3%;  Score 32.6;  DB 9;  Length 282;
  Best Local Similarity   58.9%;  Pred. No. 2.7e+02;
  Matches   56;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy       6 ccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggg 65
           ||| |     ||| ||   |   | | || || || | |||| ||||| ||||| | |||
Db     129 CCCGCCCTCCAGGCGTCCCCCAAGGACCTCATAGGGAGCCAGGGGGGCAGGGGCCGGGGG 188

Qy      66 gagtcctcctggagcctggtgccctagaagcccac 100
           ||||   ||     |   | ||| |   |||||||
Db     189 GAGTTTCCCATAGTCTACGAGCCATCCCAGCCCAC 223


RESULT   7
AW026388/c
LOCUS       AW026388                354 bp    mRNA    linear   EST 09-MAR-2000
DEFINITION  wv12f05.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:2529345 3'
            similar to contains element MSR1 repetitive element ;, mRNA
            sequence.
ACCESSION   AW026388
VERSION     AW026388.1  GI:5879918
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 354)
  AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
            Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1245   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 338.
FEATURES             Location/Qualifiers
     source          1. .354
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2529345"
                     /clone_lib="NCI_CGAP_Brn23"
                     /tissue_type="glioblastoma (pooled)"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATATCTTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library is normalized, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT       72 a    104 c     89 g     88 t      1 others
ORIGIN


  Query Match             32.3%;  Score 32.6;  DB 9;  Length 354;
  Best Local Similarity   73.2%;  Pred. No. 2.7e+02;
  Matches   41;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      20 gtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagtcctcct 75
           | |||  ||| ||||| |||  |||| | |||| ||||| || |||||| || | |
Db     213 GGGGGGGGGGGATCTTTTGGAAACCCNGGGGGGGGGGGGGGGGGGGAGCCCCCAT 158


RESULT   8
BF808916/c
LOCUS       BF808916                539 bp    mRNA    linear   EST 12-JAN-2001
DEFINITION  PM4-CI0154-151100-001-h12 CI0154 Homo sapiens cDNA, mRNA sequence.
ACCESSION   BF808916
VERSION     BF808916.1  GI:12137905
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 539)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
            ,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
  MEDLINE   20202663
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/scripts/gethtml2.pl?t1=PM4&t2=PM4-CI0154-
            151100-001-h12&t3=2000-11-15&t4=1)
            Seq primer: puc 18 forward
            High quality sequence start: 20
            High quality sequence stop: 539.
FEATURES                Location/Qualifiers
     source             1. .539
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /clone_lib="CI0154"
                        /dev_stage="Adult"
                        /note="Organ: colon_ins; Vector: puc18; Site_1: SmaI;
                        Site_2: SmaI; A mini-library was made by cloning products
                        derived from  ORESTES PCR (U.S. Letters Patent application
                        No. 196,716 -  Ludwig Institute for Cancer Research)
                        profiles into  the pUC 18 vector. Reverse transcription of
                        tissue mRNA and cDNA amplification were performed under
                        low stringency conditions."
BASE COUNT      131 a    140 c    171 g     97 t
ORIGIN


  Query Match             32.3%;  Score 32.6;  DB 10;  Length 539;
  Best Local Similarity   58.9%;  Pred. No. 2.8e+02;
  Matches   56;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy       7 cccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggagggg 66
           ||||||  | |||||||| ||||  |||   | ||  || |   ||| ||  |   |
Db     357 CCCCTGCAGGGGAGTGGGGTGGGTCTCTCGAGTGCGGCCTGCCAGGCAGGAACTTGGCTC 298

Qy      67 agtcctcctggagcctggtgccctagaagcccacg 101
           |   | | || ||||  |   ||| | |||||| |
Db     297 ATAGCCCATGCAGCCATGCAGCCTTGCAGCCCAGG 263


RESULT  9
AL585199
LOCUS       AL585199                685 bp    mRNA    linear   EST 28-FEB-2001
DEFINITION  AL585199 BP Chicken Embryo Library Gallus gallus cDNA clone
            ROS020D01, mRNA sequence.
ACCESSION   AL585199
VERSION     AL585199.1  GI:13163932
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 685)
  AUTHORS   Murray,F.
  TITLE     BP Chicken Embryo Library
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Frazer Murray
            Dept. Genomics and Bioinformatics
            Roslin Institute
            Roslin, Midlothian, EH25 9PS, UK
            Tel: +44 (0)131 527 4200
            Fax: +44 (0)131 440 0434
            Email: frazer.murray@bbsrc.ac.uk
            Seq primer: T7.
FEATURES                Location/Qualifiers
     source             1. .685
                        /organism="Gallus gallus"
                        /db_xref="taxon:9031"
                        /clone="ROS020D01"
                        /clone_lib="BP Chicken Embryo Library"
                        /tissue_type="Embryo"
                        /dev_stage="5 days old"
                        /lab_host="DH10B"
                        /note="Vector: pBLUESCRIPT SK; Site_1: Not1; Site_2: Sal1;
                        Cloned unidirectionally. Primer: Oligo dT. 5' adaptor
                        sequence: 5' TCGACCTCGAG 3'; 3' adaptor sequence: 5'
                        GCGGCCGCTTTTTTTTTTTTTTTTT 3'"
BASE COUNT      134 a    222 c    224 g    101 t      4 others
ORIGIN


  Query Match             31.9%;  Score 32.2;  DB 9;  Length 685;
  Best Local Similarity   65.7%;  Pred. No. 3.6e+02;
  Matches   46;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy      17 ggagtgggctgggaatcttatgggcacccagaggggcgggggcggagggggagtcctcctg 76
           ||||  |||  ||  |  |  ||| | | ||| ||| || |||||| ||  || ||  | |
Db     364 GGAGGAGGCGGGCGACATGGTGGCCGCGCAGGGGGACGAGGGCGGNGGCTAGCCCGGCCG 423

Qy      77 gagcctggtg 86
           |||| || ||
Db     424 GAGCTTGTTG 433


RESULT  10
BG309093
LOCUS       BG309093                735 bp    mRNA    linear   EST 22-OCT-2001
DEFINITION  HVSMEc0002B16f Hordeum vulgare seedling shoot EST library
            HVcDNA0003 (Etiolated and unstressed) Hordeum vulgare cDNA clone
            HVSMEc0002B16f, mRNA sequence.
ACCESSION   BG309093
VERSION     BG309093.2  GI:16313793
KEYWORDS    EST.
SOURCE      barley.
  ORGANISM  Hordeum vulgare
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
            ; Triticeae; Hordeum.
REFERENCE   1  (bases 1 to 735)
  AUTHORS   Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D., Yu
            ,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi,D.W., Fenton
            ,R.D., Oates,R. and Main,D.
  TITLE     Development of a genetically and physically anchored EST resource
            for barley genomics: Morex unstressed seedling shoot cDNA library
  JOURNAL   Unpublished (2001)
COMMENT     On Feb 22, 2001 this sequence version replaced gi:13109940.
            Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Total hq bases = 300
            Seq primer: AATTAACCCTCACTAAAGGG
            High quality sequence stop: 718.
FEATURES                Location/Qualifiers
     source             1. .735
```

/organism="Hordeum vulgare"
/cultivar="Morex"
/db_xref="taxon:4513"
/clone="HVSMEc0002B16f"
/clone_lib="Hordeum vulgare seedling shoot EST library HVcDNA0003 (Etiolated and unstressed)"
/tissue_type="Seedling shoot"
/lab_host="TJC121"
/note="Vector: lambdaZAP; Site_1: EcoR1; Site_2: Xho1; Seeds were surface sterilized then germinated under axenic conditions in the dark at room temperature on filter paper with water, nystatin and cefotaxime in covered crystallization dishes. Five-day old seedling shoots were then harvested, total RNA was prepared, poly(A) RNA was purified, one primary unamplified cDNA library was made, and 1 million pfu were in vivo excised to give pBluescript SK(-) cDNA phagemids. These steps were performed in the TJ Close laboratory at the University of California, Riverside (Choi, Close, Fenton). Phagemids were plated and picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were performed at CUGI (Wing, Yu, Frisch, Henry, Simmons, Oates , Rambo, Main). The sequence has been trimmed to remove vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see http://www.genome.clemson.edu/projects/barley. To order this clone see http://www.genome.clemson.edu/orders Also see Close TJ, Wing R, Kleinhofs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"
BASE COUNT 189 a 156 c 264 g 116 t 10 others
ORIGIN

Query Match 31.7%; Score 32; DB 10; Length 735;
Best Local Similarity 62.5%; Pred. No. 4e+02;
Matches 50; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

```
Qy      10 ctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagt 69
           |||||| | || |||  ||| |      ||| |||||  |||||| | ||||| ||||||
Db     559 CTGGTGGGCAGGGGGTGGGGGAGGGGGAGGGAACCCACCGGGGCGTGAGCGGAAGGGAGT 618

Qy      70 cctcctggagcctggtgccc 89
               |||      | ||||
Db     619 TTGGGTGGGTGTGTGGGCCC 638
```

RESULT 11
BE600720/c
LOCUS BE600720 382 bp mRNA linear EST 18-AUG-2000
DEFINITION PI1_91_G10.b1_A002 Pathogen induced 1 (PI1) Sorghum bicolor cDNA, mRNA sequence.
ACCESSION BE600720
VERSION BE600720.1 GI:9855905
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 382)
AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt ,L.H.
TITLE An EST database from Sorghum: pathogen-induced plants
JOURNAL Unpublished (2000)
COMMENT Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.
Seq primer: JEN REV
High quality sequence stop: 382
POLYA=No.
FEATURES Location/Qualifiers
source 1. .382
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Pathogen induced 1 (PI1)"
/note="Organ: Anthracnose-infected leaves from two-week-old sorghum plants 48 hr after inoculation; Vector: pBluescript II from Lambda Zap II; Site_1: XhoI; Site_2: EcoRI; Two-week-old sorghum plants (BTX 623 cultivar) were infected with pathogen (isolate FRM42I of Colletotrichum graminicola, which is a sorghum isolate). RNA was prepared from infected leaves harvested from 45 seedlings 48 hours after inoculation. Note: young seedlings (2 weeks old) exhibit juvenile resistant reaction, which is an incompatible interaction. As they grow older (4 weeks or older), plants resume susceptiblity to anthracnose disease. The library was made from poly-A RNA in the cloning vector lambda ZAP II. Clones to be sequenced were prepared by mass excision. WARNING: While most or all ESTs are expected to derive from the host plant, no effort was made to eliminate ESTs deriving from the pathogen."
BASE COUNT 67 a 134 c 90 g 91 t
ORIGIN

Query Match 31.5%; Score 31.8; DB 10; Length 382;
Best Local Similarity 59.3%; Pred. No. 4.2e+02;
Matches 54; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```
Qy      11 tggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagtc 70
           || ||| | |||||   ||   || ||| |||  |  ||| || |  | ||||| |||
Db     103 TGATGATGTGTGGGGGTGGTGGCTCATGCGCAGGCGCAGGCGCAGCAGAGGAGGCGAGAG 44

Qy      71 ctcctggagcctggtgccctagaagcccacg 101
               ||  |||||| ||   | ||||| |
Db      43 GGAGAAGATGCTGGTGTCCGCCATGCCCATG 13
```

RESULT 12
BE600139/c
LOCUS BE600139 463 bp mRNA linear EST 18-AUG-2000
DEFINITION PI1_79_G07.b1_A002 Pathogen induced 1 (PI1) Sorghum bicolor cDNA, mRNA sequence.
ACCESSION BE600139
VERSION BE600139.1 GI:9855212
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 463)
AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt ,L.H.
TITLE An EST database from Sorghum: pathogen-induced plants
JOURNAL Unpublished (2000)
COMMENT Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions

below Phred quality 16. The threshold for highest quality sequence is 20.
Seq primer: JEN REV
High quality sequence stop: 462
POLYA=No.

FEATURES Location/Qualifiers
source 1. .463
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Pathogen induced 1 (PI1)"
/note="Organ: Anthracnose-infected leaves from two-week-old sorghum plants 48 hr after inoculation; Vector: pBluescript II from Lambda Zap II; Site_1: XhoI; Site_2: EcoRI; Two-week-old sorghum plants (BTX 623 cultivar) were infected with pathogen (isolate FRM42I of Colletotrichum graminicola, which is a sorghum isolate). RNA was prepared from infected leaves harvested from 45 seedlings 48 hours after inoculation. Note: young seedlings (2 weeks old) exhibit juvenile resistant reaction, which is an incompatible interaction. As they grow older (4 weeks or older), plants resume susceptiblity to anthracnose disease. The library was made from poly-A RNA in the cloning vector lambda ZAP II. Clones to be sequenced were prepared by mass excision. WARNING: While most or all ESTs are expected to derive from the host plant, no effort was made to eliminate ESTs deriving from the pathogen."

BASE COUNT 79 a 154 c 124 g 106 t
ORIGIN

Query Match 31.5%; Score 31.8; DB 10; Length 463;
Best Local Similarity 59.3%; Pred. No. 4.3e+02;
Matches 54; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```
Qy      11 tggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggagggagtc 70
           || ||| | |||||   ||   || ||| |||  |  ||| || |  | ||||| |||
Db     114 TGATGATGTGTGGGGGTGGTGGCTCATGCGCAGGCGCAGGCGCAGCAGAGGAGGCGAGAG 55

Qy      71 ctcctggagcctggtgccctagaagcccacg 101
              ||  |||||| ||   | |||||
Db      54 GGAGAAGATGCTGGTGTCCGCCATGCCCATG 24
```

RESULT 13
BE598197/c
LOCUS BE598197 474 bp mRNA linear EST 18-AUG-2000
DEFINITION PI1_66_H03.b1_A002 Pathogen induced 1 (PI1) Sorghum bicolor cDNA, mRNA sequence.
ACCESSION BE598197
VERSION BE598197.1 GI:9853270
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 474)
AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt ,L.H.
TITLE An EST database from Sorghum: pathogen-induced plants
JOURNAL Unpublished (2000)
COMMENT Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.
Seq primer: JEN REV
High quality sequence stop: 416
POLYA=No.

FEATURES Location/Qualifiers
source 1. .474
/organism="Sorghum bicolor"
/db_xref="taxon:4558"
/clone_lib="Pathogen induced 1 (PI1)"
/note="Organ: Anthracnose-infected leaves from two-week-old sorghum plants 48 hr after inoculation; Vector: pBluescript II from Lambda Zap II; Site_1: XhoI; Site_2: EcoRI; Two-week-old sorghum plants (BTX 623 cultivar) were infected with pathogen (isolate FRM42I of Colletotrichum graminicola, which is a sorghum isolate). RNA was prepared from infected leaves harvested from 45 seedlings 48 hours after inoculation. Note: young seedlings (2 weeks old) exhibit juvenile resistant reaction, which is an incompatible interaction. As they grow older (4 weeks or older), plants resume susceptiblity to anthracnose disease. The library was made from poly-A RNA in the cloning vector lambda ZAP II. Clones to be sequenced were prepared by mass excision. WARNING: While most or all ESTs are expected to derive from the host plant, no effort was made to eliminate ESTs deriving from the pathogen."

BASE COUNT 82 a 160 c 126 g 106 t
ORIGIN

Query Match 31.5%; Score 31.8; DB 10; Length 474;
Best Local Similarity 59.3%; Pred. No. 4.3e+02;
Matches 54; Conservative 0; Mismatches 37; Indels 0; Gaps 0;

```
Qy      11 tggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggagggagtc 70
           || ||| | |||||   ||   || ||| |||  |  ||| || |  | ||||| |||
Db     105 TGATGATGTGTGGGGGTGGTGGCTCATGCGCAGGCGCAGGCGCAGCAGAGGAGGCGAGAG 46

Qy      71 ctcctggagcctggtgccctagaagcccacg 101
              ||  |||||| ||   | |||||
Db      45 GGAGAAGATGCTGGTGTCCGCCATGCCCATG 15
```

RESULT 14
BE600794/c
LOCUS BE600794 487 bp mRNA linear EST 18-AUG-2000
DEFINITION PI1_90_C08.b1_A002 Pathogen induced 1 (PI1) Sorghum bicolor cDNA, mRNA sequence.
ACCESSION BE600794
VERSION BE600794.1 GI:9855769
KEYWORDS EST.
SOURCE sorghum.
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE 1 (bases 1 to 487)
AUTHORS Cordonnier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt ,L.H.
TITLE An EST database from Sorghum: pathogen-induced plants
JOURNAL Unpublished (2000)
COMMENT Contact: Cordonnier-Pratt MM
Department of Botany
The University of Georgia
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 542 1805
Email: mmpratt@uga.edu
Sequences have been trimmed to exclude PolyA, vector and regions below Phred quality 16. The threshold for highest quality sequence is 20.
Seq primer: JEN REV
High quality sequence stop: 486
POLYA=No.

FEATURES Location/Qualifiers

```
     source          1. .487
                     /organism="Sorghum bicolor"
                     /db_xref="taxon:4558"
                     /clone_lib="Pathogen induced 1 (PI1)"
                     /note="Organ: Anthracnose-infected leaves from
                     two-week-old sorghum plants 48 hr after inoculation;
                     Vector: pBluescript II from Lambda Zap II; Site_1: XhoI;
                     Site_2: EcoRI; Two-week-old sorghum plants (BTX 623
                     cultivar) were infected with pathogen (isolate FRM42I of
                     Colletotrichum graminicola, which is a sorghum isolate).
                     RNA was prepared from infected leaves harvested from 45
                     seedlings  48 hours after inoculation.  Note: young
                     seedlings (2 weeks old) exhibit juvenile resistant
                     reaction, which is an incompatible interaction. As they
                     grow older (4 weeks or older), plants resume susceptiblity
                     to anthracnose disease. The library was made from poly-A
                     RNA in the cloning vector lambda ZAP II. Clones to be
                     sequenced were prepared by mass excision. WARNING: While
                     most or all ESTs are expected to derive from the host
                     plant, no effort was made to eliminate ESTs deriving from
                     the pathogen."
BASE COUNT       84 a    163 c    132 g    108 t
ORIGIN


  Query Match             31.5%;  Score 31.8;  DB 10;  Length 487;
  Best Local Similarity   59.3%;  Pred. No. 4.3e+02;
  Matches   54;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy      11 tggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagtc 70
           || ||| | ||||| ||  || ||| ||| | ||| || | | ||||| |||
Db     114 TGATGATGTGTGGGGGTGGTGGCTCATGCGCAGGCGCAGGCGCAGCAGAGGAGGCGAGAG 55

Qy      71 ctcctggagcctggtgccctagaagcccacg 101
              ||  |||||| ||   | ||||| |
Db      54 GGAGAAGATGCTGGTGTCCGCCATGCCCATG 24


RESULT  15
BE595429/c
LOCUS       BE595429                 532 bp    mRNA    linear   EST 18-AUG-2000
DEFINITION  PI1_55_D02.b1_A002 Pathogen induced 1 (PI1) Sorghum bicolor cDNA,
            mRNA sequence.
ACCESSION   BE595429
VERSION     BE595429.1  GI:9850502
KEYWORDS    EST.
SOURCE      sorghum.
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 532)
  AUTHORS   Cordonnier-Pratt,M.-M., Gingle,A., Dean,R., Sudman,M. and Pratt
            ,L.H.
  TITLE     An EST database from Sorghum: pathogen-induced plants
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Cordonnier-Pratt MM
            Department of Botany
            The University of Georgia
            Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
            Tel: 706 542 1860
            Fax: 706 542 1805
            Email: mmpratt@uga.edu
            Sequences have been trimmed to exclude PolyA, vector and regions
            below Phred quality 16. The threshold for highest quality sequence
            is 20.
            Seq primer: JEN REV
            High quality sequence stop: 525
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .532
                     /organism="Sorghum bicolor"
                     /db_xref="taxon:4558"
                     /clone_lib="Pathogen induced 1 (PI1)"
                     /note="Organ: Anthracnose-infected leaves from
                     two-week-old sorghum plants 48 hr after inoculation;
                     Vector: pBluescript II from Lambda Zap II; Site_1: XhoI;
                     Site_2: EcoRI; Two-week-old sorghum plants (BTX 623
                     cultivar) were infected with pathogen (isolate FRM42I of
                     Colletotrichum graminicola, which is a sorghum isolate).
                     RNA was prepared from infected leaves harvested from 45
                     seedlings  48 hours after inoculation.  Note: young
                     seedlings (2 weeks old) exhibit juvenile resistant
                     reaction, which is an incompatible interaction. As they
                     grow older (4 weeks or older), plants resume susceptiblity
                     to anthracnose disease. The library was made from poly-A
                     RNA in the cloning vector lambda ZAP II. Clones to be
                     sequenced were prepared by mass excision. WARNING: While
                     most or all ESTs are expected to derive from the host
                     plant, no effort was made to eliminate ESTs deriving from
                     the pathogen."
BASE COUNT       96 a    175 c    145 g    116 t
ORIGIN


  Query Match             31.5%;  Score 31.8;  DB 10;  Length 532;
  Best Local Similarity   59.3%;  Pred. No. 4.4e+02;
  Matches   54;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy      11 tggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagtc 70
           || ||| | ||||| ||  || ||| ||| | ||| || | | ||||| |||
Db     103 TGATGATGTGTGGGGGTGGTGGCTCATGCGCAGGCGCAGGCGCAGCAGAGGAGGCGAGAG 44

Qy      71 ctcctggagcctggtgccctagaagcccacg 101
              ||  |||||| ||   | ||||| |
Db      43 GGAGAAGATGCTGGTGTCCGCCATGCCCATG 13


Search completed: October  3, 2002, 16:14:49
Job time: 16944 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 16:18:58 ; Search time 180.77 Seconds
(without alignments)
137.241 Million cell updates/sec

Title: US-09-826-581-3_COPY_600_700
Perfect score: 101
Sequence: 1 aggtccccccctggtgaggag.....tggtgccctagaagcccacg 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	27.8	27.5	722	1	US-08-702-344-8	Sequence 8, Appli
2	27.8	27.5	4411529	4	US-09-103-840A-1	Sequence 1, Appli
3	27.6	27.3	6822	4	US-09-426-998-3	Sequence 3, Appli
4	27.6	27.3	7741	4	US-09-426-998-4	Sequence 4, Appli
5	27.2	26.9	31571	1	US-08-323-443B-1	Sequence 1, Appli
6	27.2	26.9	53526	3	US-08-658-136-2	Sequence 2, Appli
7	27.2	26.9	53577	3	US-08-658-136-1	Sequence 1, Appli
c 8	26.8	26.5	1190	1	US-08-451-947-9	Sequence 9, Appli
c 9	26.8	26.5	1190	2	US-08-424-826A-9	Sequence 9, Appli
c 10	26.8	26.5	1190	3	US-08-928-694-9	Sequence 9, Appli
c 11	26.8	26.5	1190	5	PCT-US91-06950-9	Sequence 9, Appli
c 12	26.8	26.5	3286	4	US-09-211-417-2	Sequence 2, Appli
13	26.8	26.5	4183	3	US-08-996-083-2	Sequence 2, Appli
14	26.8	26.5	4183	4	US-09-429-516-2	Sequence 2, Appli
c 15	26.6	26.3	573	2	US-08-290-665A-140	Sequence 140, App
c 16	26.6	26.3	573	5	PCT-US95-10398-140	Sequence 140, App
c 17	26.4	26.1	6763	2	US-08-756-506-23	Sequence 23, Appl
c 18	26.4	26.1	10807	1	US-08-206-176-7	Sequence 7, Appli
c 19	26.4	26.1	10807	2	US-08-756-506-5	Sequence 5, Appli
c 20	26.2	25.9	573	2	US-08-290-665A-139	Sequence 139, App
c 21	26.2	25.9	573	5	PCT-US95-10398-139	Sequence 139, App
c 22	26.2	25.9	957	4	US-08-836-075A-17	Sequence 17, Appl
23	26.2	25.9	4403765	4	US-09-103-840A-2	Sequence 2, Appli
c 24	26	25.7	775	4	US-08-998-416-580	Sequence 580, App
c 25	25.8	25.5	462	3	US-08-718-904-118	Sequence 118, App
c 26	25.8	25.5	573	2	US-08-290-665A-144	Sequence 144, App
c 27	25.8	25.5	573	5	PCT-US95-10398-144	Sequence 144, App
c 28	25.8	25.5	1875	4	US-09-422-869-21	Sequence 21, Appl
c 29	25.8	25.5	3247	4	US-08-718-388-4	Sequence 4, Appli
c 30	25.8	25.5	7824	4	US-08-718-388-6	Sequence 6, Appli
c 31	25.8	25.5	16382	4	US-08-718-388-8	Sequence 8, Appli
c 32	25.8	25.5	49136	4	US-09-422-869-1	Sequence 1, Appli
33	25.6	25.3	603	4	US-09-134-246-7	Sequence 7, Appli
c 34	25.6	25.3	685	1	US-08-451-947-7	Sequence 7, Appli
c 35	25.6	25.3	685	2	US-08-424-826A-7	Sequence 7, Appli
c 36	25.6	25.3	685	3	US-08-928-694-7	Sequence 7, Appli
c 37	25.6	25.3	685	5	PCT-US91-06950-7	Sequence 7, Appli
38	25.6	25.3	11958	4	US-09-134-246-8	Sequence 8, Appli
c 39	25.6	25.3	4403765	4	US-09-103-840A-2	Sequence 2, Appli
c 40	25.6	25.3	4411529	4	US-09-103-840A-1	Sequence 1, Appli
41	25.4	25.1	1960	2	US-08-533-306A-1	Sequence 1, Appli
42	25.4	25.1	1960	2	US-08-742-923A-1	Sequence 1, Appli
43	25.4	25.1	2680	2	US-08-533-306A-5	Sequence 5, Appli
44	25.4	25.1	2680	2	US-08-742-923A-5	Sequence 5, Appli
45	25.4	25.1	2887	2	US-08-533-306A-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-702-344-8
; Sequence 8, Application US/08702344
; Patent No. 5723315
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John
; APPLICANT: LaVallie, Edward
; APPLICANT: Racie, Lisa
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/702,344
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Brown, Scott A.
; REGISTRATION NUMBER: 32,724
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8224
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 722 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-702-344-8

Query Match 27.5%; Score 27.8; DB 1; Length 722;
Best Local Similarity 62.0%; Pred. No. 10;
Matches 44; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

```
Qy      20 gtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagtcctcctggag 79
           ||| ||| |||| | || |    || || | || |   | |||| || ||| |
Db      38 GTGCCCTGAGAATGCTGAGGACTGTGAGCAGGCTGTGGTCCACGTGGAGACCACCTTGTA 97

Qy      80 cctggtgccct 90
           ||||||||||||
Db      98 CCTGGTGCCCT 108


RESULT   2
US-09-103-840A-1
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
;   LENGTH: 4411529
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   OTHER INFORMATION: H37Rv
US-09-103-840A-1


  Query Match           27.5%;  Score 27.8;  DB 4;  Length 4411529;
  Best Local Similarity   55.8%;  Pred. No. 21;
  Matches   53;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy       1 aggtccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcg 60
           ||| ||| ||  |||| ||  || ||  | | |      |||||  || |  ||| ||
Db   11849 aggccccaccgtgtgacgaccggcctcaggagcgcggttgcacctcgacgctcggtggtc 11908

Qy      61 gaggggagtcctcctggagcctggtgccctagaag 95
           | || ||   |||| |  ||   | || |  || |
Db   11909 ggggcgacggctccggtcgcgacgagcgccggacg 11943


RESULT   3
US-09-426-998-3
; Sequence 3, Application US/09426998
; Patent No. 6358706
; GENERAL INFORMATION:
;  APPLICANT: DUBIN, ADRIENNE E.
;  APPLICANT:  PYATI, JAYASHREE
;  APPLICANT:  ZHU, JESSICA Y
;  APPLICANT:  ERLANDER, MARK G
;  APPLICANT:  GALINDO, JOSE E
;  TITLE OF INVENTION: DNA ENCODING HUMAN ALPHA1G T-TYPE CALCIUM
;  TITLE OF INVENTION:  CHANNEL (ALPHA1G-C)
;  FILE REFERENCE: ORT-1057
;  CURRENT APPLICATION NUMBER: US/09/426,998
;  CURRENT FILING DATE:  1999-10-26
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PATENTIN VER. 2.0
; SEQ ID NO 3
;   LENGTH: 6822
;   TYPE: DNA
;   ORGANISM: HOMO SAPIENS
US-09-426-998-3


  Query Match           27.3%;  Score 27.6;  DB 4;  Length 6822;
  Best Local Similarity   56.7%;  Pred. No. 14;
  Matches   51;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy       3 gtccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcgga 62
           ||| || | |||||  | |||  ||| | | | | || |||||   ||| | | | |
Db    5359 gtctccacaggtgacaattggaatggcattatgaaggacaccctccgggactgtgaccag 5418

Qy      63 ggggagtcctcctggagcctggtgccctag 92
           | |    ||| ||  | |  |||   || |
Db    5419 gagtccacctgctacaacacggtcatctcg 5448


RESULT   4
US-09-426-998-4
; Sequence 4, Application US/09426998
; Patent No. 6358706
; GENERAL INFORMATION:
;  APPLICANT: DUBIN, ADRIENNE E.
;  APPLICANT:  PYATI, JAYASHREE
;  APPLICANT:  ZHU, JESSICA Y
;  APPLICANT:  ERLANDER, MARK G
;  APPLICANT:  GALINDO, JOSE E
;  TITLE OF INVENTION: DNA ENCODING HUMAN ALPHA1G T-TYPE CALCIUM
;  TITLE OF INVENTION:  CHANNEL (ALPHA1G-C)
;  FILE REFERENCE: ORT-1057
;  CURRENT APPLICATION NUMBER: US/09/426,998
;  CURRENT FILING DATE:  1999-10-26
;  NUMBER OF SEQ ID NOS: 5
;  SOFTWARE: PATENTIN VER. 2.0
; SEQ ID NO 4
;   LENGTH: 7741
;   TYPE: DNA
;   ORGANISM: HOMO SAPIENS
US-09-426-998-4


  Query Match           27.3%;  Score 27.6;  DB 4;  Length 7741;
  Best Local Similarity   56.7%;  Pred. No. 15;
  Matches   51;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy       3 gtccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcgga 62
           ||| || | |||||  | |||  ||| | | | | || |||||   ||| | | | |
Db    5881 gtctccacaggtgacaattggaatggcattatgaaggacaccctccgggactgtgaccag 5940

Qy      63 ggggagtcctcctggagcctggtgccctag 92
           | |    ||| ||  | |  |||   || |
Db    5941 gagtccacctgctacaacacggtcatctcg 5970


RESULT   5
US-08-323-443B-1
; Sequence 1, Application US/08323443B
; Patent No. 5654170
;  GENERAL INFORMATION:
;    APPLICANT:  KLINGER, KATHERINE W.
;    APPLICANT:  LANDES, GREGORY M.
;    APPLICANT:  BURN, TIMOTHY C.
;    APPLICANT:  CONNORS, TIMOTHY D.
;    APPLICANT:  DACKOWSKI, WILLIAM R.
;    APPLICANT:  GERMINO, GREGORY
;    APPLICANT:  QIAN, FENG
;    TITLE OF INVENTION:  POLYCYSTIC KIDNEY DISEASE GENE
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Darby & Darby PC
;      STREET:  805 Third Avenue
;      CITY:  New York
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/323,443B
;       FILING DATE:  12-OCT-1994
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Ludwig, S. Peter
;       REGISTRATION NUMBER:  25,351
;       REFERENCE/DOCKET NUMBER:  0372/0A462
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 527-7700
;       TELEFAX:  (212) 753-6237
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  31571 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ORIGINAL SOURCE:
;       ORGANISM:  Homo sapiens
;     IMMEDIATE SOURCE:
;       CLONE:  PKD1 GENOMIC
US-08-323-443B-1


  Query Match             26.9%;  Score 27.2;  DB 1;  Length 31571;
  Best Local Similarity   58.8%;  Pred. No. 22;
  Matches   47;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy       5 ccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggagg 64
           ||  |  |||  |   |||||  |||   |   ||||  |   |||||  ||  ||  |||||
Db   16428 CCTGCTGGGTCACTCGTGGGGAGGGTGACACCTGGGGAAGTAGAGGCCCGTGGCAGGAGG 16487

Qy      65 ggagtcctcctggagcctgg 84
             |||  ||||  ||  |  ||
Db   16488 TGAGGCCTCGGGGTCCTGGG 16507


RESULT   6
US-08-658-136-2
; Sequence 2, Application US/08658136
; Patent No. 6071717
;  GENERAL INFORMATION:
;     APPLICANT:  KLINGER, KATHERINE W
;     APPLICANT:  LANDES, GREGORY M
;     APPLICANT:  BURN, TIMOTHY C
;     APPLICANT:  CONNORS, TIMOTHY D
;     APPLICANT:  DACKOWSKI, WILLIAM
;     APPLICANT:  GERMINO, GREGORY
;     APPLICANT:  QIAN, FENG
;     TITLE OF INVENTION:  POLYCYSTIC KIDNEY DISEASE GENE
;     NUMBER OF SEQUENCES:  58
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  GENZYME CORPORATION
;       STREET:  ONE MOUNTAIN ROAD
;       CITY:  FRAMINGHAM
;       STATE:  MASSACHUSETTS
;       COUNTRY:  USA
;       ZIP:  01701
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/658,136
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  LASSEN, ELIZABETH
;       REGISTRATION NUMBER:  31,845
;       REFERENCE/DOCKET NUMBER:  GEN4-17.8
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  508-872-8400
;       TELEFAX:  508-872-5415
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  53526 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-08-658-136-2


  Query Match             26.9%;  Score 27.2;  DB 3;  Length 53526;
  Best Local Similarity   58.8%;  Pred. No. 23;
  Matches   47;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy       5 ccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggagg 64
           ||  |  |||  |   |||||  |||   |   ||||  |   |||||  ||  ||  |||||
Db   17414 CCTGCTGGGTCACTCGTGGGGAGGGTGACACCTGGGGAAGTAGAGGCCCGTGGCAGGAGG 17473

Qy      65 ggagtcctcctggagcctgg 84
             |||  ||||  ||  |  ||
Db   17474 TGAGGCCTCGGGGTCCTGGG 17493


RESULT   7
US-08-658-136-1
; Sequence 1, Application US/08658136
; Patent No. 6071717
;  GENERAL INFORMATION:
;     APPLICANT:  KLINGER, KATHERINE W
;     APPLICANT:  LANDES, GREGORY M
;     APPLICANT:  BURN, TIMOTHY C
;     APPLICANT:  CONNORS, TIMOTHY D
;     APPLICANT:  DACKOWSKI, WILLIAM
;     APPLICANT:  GERMINO, GREGORY
;     APPLICANT:  QIAN, FENG
;     TITLE OF INVENTION:  POLYCYSTIC KIDNEY DISEASE GENE
;     NUMBER OF SEQUENCES:  58
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  GENZYME CORPORATION
;       STREET:  ONE MOUNTAIN ROAD
;       CITY:  FRAMINGHAM
;       STATE:  MASSACHUSETTS
;       COUNTRY:  USA
;       ZIP:  01701
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/658,136
;       FILING DATE:
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  LASSEN, ELIZABETH
;       REGISTRATION NUMBER:  31,845
;       REFERENCE/DOCKET NUMBER:  GEN4-17.8
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  508-872-8400
;       TELEFAX:  508-872-5415
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  53577 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-08-658-136-1
```

```
  Query Match             26.9%;  Score 27.2;  DB 3;  Length 53577;
  Best Local Similarity   58.8%;  Pred. No. 23;
  Matches   47;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy       5 ccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggagg 64
           ||  |  ||| |   |||||  |||   |   |||| |   |||||  || ||  |||||
Db   17413 CCTGCTGGGTCACTCGTGGGGAGGGTGACACCTGGGGAAGTAGAGGCCCGTGGCAGGAGG 17472

Qy      65 ggagtcctcctggagcctgg 84
             ||| ||||  ||  |  ||
Db   17473 TGAGGCCTCGGGGTCCTGGG 17492


RESULT   8
US-08-451-947-9/c
; Sequence 9, Application US/08451947
; Patent No. 5702906
;  GENERAL INFORMATION:
;    APPLICANT:  GENENTECH, INC.
;    APPLICANT:  ROSENTHAL, ARNON
;    TITLE OF INVENTION:  NOVEL NEUROTROPHIC FACTOR
;    NUMBER OF SEQUENCES:  100
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/451,947
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/426419
;      FILING DATE:  19-APR-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/030013
;      FILING DATE:  22-MAR-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/648482
;      FILING DATE:  31-JAN
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/587707
;      FILING DATE:  1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Torchia, Timothy E.
;      REGISTRATION NUMBER:  36,700
;      REFERENCE/DOCKET NUMBER:  666P2C1D2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-8674
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1190 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-451-947-9


  Query Match             26.5%;  Score 26.8;  DB 1;  Length 1190;
  Best Local Similarity   57.0%;  Pred. No. 20;
  Matches   49;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy      11 tggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagtc 70
             ||  |||||  ||| | ||| ||  |||      |  | | || ||     ||||||  |
Db     469 TGTCGAGGATGGGGTTAGGACTCCAATGACACACTGGGGAGGAGGAAAATGAGGGGGATG 410

Qy      71 ctcctggagcctggtgccctagaagc 96
             |    |||||||||  |    || |||
Db     409 CGGAGGGAGCCTGGGGGAGCAGGAGC 384


RESULT   9
US-08-424-826A-9/c
; Sequence 9, Application US/08424826A
; Patent No. 5830858
;  GENERAL INFORMATION:
;    APPLICANT:  Rosenthal, Arnon
;    TITLE OF INVENTION:  NOVEL NEUROTROPHIC FACTOR
;    NUMBER OF SEQUENCES:  98
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/424,826A
;      FILING DATE:  19-Apr-1995
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/240387
;      FILING DATE:  10-May-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/648482
;      FILING DATE:  31-JAN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/587707
;      FILING DATE:  25-SEP-1990
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Torchia, PhD., Timothy E.
;      REGISTRATION NUMBER:  36,700
;      REFERENCE/DOCKET NUMBER:  P0666P1C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-8674
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  9:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1190 base pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
US-08-424-826A-9


  Query Match             26.5%;  Score 26.8;  DB 2;  Length 1190;
  Best Local Similarity   57.0%;  Pred. No. 20;
  Matches   49;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy      11 tggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagtc 70
             ||  |||||  ||| | ||| ||  |||      |  | | || ||     ||||||  |
Db     469 TGTCGAGGATGGGGTTAGGACTCCAATGACACACTGGGGAGGAGGAAAATGAGGGGGATG 410

Qy      71 ctcctggagcctggtgccctagaagc 96
             |    |||||||||  |    || |||
Db     409 CGGAGGGAGCCTGGGGGAGCAGGAGC 384


RESULT  10
US-08-928-694-9/c
```

```
; Sequence 9, Application US/08928694
; Patent No. 6037320
;  GENERAL INFORMATION:
;     APPLICANT:  ROSENTHAL, ARNON
;     TITLE OF INVENTION:  NOVEL NEUROTROPHIC FACTOR
;     NUMBER OF SEQUENCES:  100
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Genentech, Inc.
;       STREET:  1 DNA Way
;       CITY:  South San Francisco
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94080
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  WinPatin (Genentech)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/928,694
;       FILING DATE:  12-Sep-1997
;       CLASSIFICATION:  424
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/451947
;       FILING DATE:  26-MAY-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/426419
;       FILING DATE:  19-APR-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/030013
;       FILING DATE:  22-MAR-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/648482
;       FILING DATE:  31-JAN
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/587707
;       FILING DATE:  1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Torchia, PhD., Timothy E.
;       REGISTRATION NUMBER:  36,700
;       REFERENCE/DOCKET NUMBER:  P0666P2C1D2C1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  650/225-8674
;       TELEFAX:  650/952-9881
;   INFORMATION FOR SEQ ID NO:  9:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1190 base pairs
;       TYPE:  Nucleic Acid
;       STRANDEDNESS:  Single
;       TOPOLOGY:  Linear
US-08-928-694-9


  Query Match             26.5%;  Score 26.8;  DB 3;  Length 1190;
  Best Local Similarity   57.0%;  Pred. No. 20;
  Matches   49;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy      11 tggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagtc 70
           ||  |||||  ||| | ||| ||  |||     |  | | || ||     ||||||  |
Db     469 TGTCGAGGATGGGGTTAGGACTCCAATGACACACTGGGGAGGAGGAAAATGAGGGGGATG 410

Qy      71 ctcctggagcctggtgccctagaagc 96
           |    ||||||||| |    || |||
Db     409 CGGAGGGAGCCTGGGGGAGCAGGAGC 384


RESULT  11
PCT-US91-06950-9/c
; Sequence 9, Application PC/TUS9106950
;  GENERAL INFORMATION:
;     APPLICANT:  GENENTECH, INC.
;     APPLICANT:  ROSENTHAL, ARNON
;     TITLE OF INVENTION:  NOVEL NEUROTROPHIC FACTOR
;     NUMBER OF SEQUENCES:  100
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Genentech, Inc.
;       STREET:  460 Point San Bruno Blvd
;       CITY:  South San Francisco
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94080
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  patin (Genentech)
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US91/06950
;       FILING DATE:  19910924
;       CLASSIFICATION:  436
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/648482
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/587707
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Hensley, Max D.
;       REGISTRATION NUMBER:  27,043
;       REFERENCE/DOCKET NUMBER:  666P1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  415/266-1994
;       TELEFAX:  415/952-9881
;       TELEX:  910/371-7168
;   INFORMATION FOR SEQ ID NO:  9:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1190 bases
;       TYPE:  NUCLEIC ACID
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
PCT-US91-06950-9


  Query Match             26.5%;  Score 26.8;  DB 5;  Length 1190;
  Best Local Similarity   57.0%;  Pred. No. 20;
  Matches   49;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy      11 tggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggaggggagtc 70
           ||  |||||  ||| | ||| ||  |||     |  | | || ||     ||||||  |
Db     469 TGTCGAGGATGGGGTTAGGACTCCAATGACACACTGGGGAGGAGGAAAATGAGGGGGATG 410

Qy      71 ctcctggagcctggtgccctagaagc 96
           |    ||||||||| |    || |||
Db     409 CGGAGGGAGCCTGGGGGAGCAGGAGC 384


RESULT  12
US-09-211-417-2/c
; Sequence 2, Application US/09211417A
; Patent No. 6177254
; GENERAL INFORMATION:
;  APPLICANT: Rattner, Jerome B
;  APPLICANT: Whitehead, Clark M
;  TITLE OF INVENTION: NUCLEOLUS AUTOANTIGENIC MARKER FOR SYSTEMIC LUPUS
;  TITLE OF INVENTION: ERHYEMATOSUS
;  TITLE OF INVENTION: GenBank
;  TITLE OF INVENTION: Accession Number U86751
;  FILE REFERENCE: UCC1
;  CURRENT APPLICATION NUMBER: US/09/211,417A
;  CURRENT FILING DATE: 1998-12-15
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 3286
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Nucleic Acid Sequence of ASE-1
```

```
; Patent No. 6177254
US-09-211-417-2


  Query Match             26.5%;  Score 26.8;  DB 4;  Length 3286;
  Best Local Similarity   68.5%;  Pred. No. 22;
  Matches   37;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

Qy      29 gaatcttatgggcacccagaggggcgggggcggaggggagtcctcctggagcct 82
           | ||| |  |||  ||| | ||||| | |||  ||| |||||| |||   ||||
Db     840 GGATCCTTAGGGTGCCCTGGGGGGCTGAGGCACAGGTGAGTCCACCTCCTGCCT 787


RESULT  13
US-08-996-083-2
; Sequence 2, Application US/08996083A
; Patent No. 6124095
; GENERAL INFORMATION:
;  APPLICANT: Magna, Holly
;  APPLICANT: Schaffer, Paul
;  APPLICANT: Lawton, Michael
;  APPLICANT: Yocum, Sue
;  APPLICANT: Mitchell, Peter
;  APPLICANT: Hutchinson, Nancy
;  APPLICANT: Murry, Lynn E.
;  TITLE OF INVENTION: HUMAN NUCLEOTIDE PYROPHOSPHOHYDROLASE-2
;  FILE REFERENCE: PF-0420 US
;  CURRENT APPLICATION NUMBER: US/08/996,083A
;  CURRENT FILING DATE: 1997-12-22
;  NUMBER OF SEQ ID NOS: 3
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 4183
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: misc_feature
;   OTHER INFORMATION: Incyte Clone No. 6124095: 1388013
;  PUBLICATION INFORMATION:
US-08-996-083-2


  Query Match             26.5%;  Score 26.8;  DB 3;  Length 4183;
  Best Local Similarity   57.0%;  Pred. No. 23;
  Matches   49;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy       5 ccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggagg 64
           |||| |||| || | |||  | |||  || | |||   | |||| | ||| | |     |
Db     576 ccccgctggggatgcgtgtcccgggcgtcctctggaggcgcagaagtgcgtgcggcctcg 635

Qy      65 ggagtcctcctggagcctggtgccct 90
           |    |    || ||||| |   |||
Db     636 gtgtccagggtgcagccttgacacct 661


RESULT  14
US-09-429-516-2
; Sequence 2, Application US/09429516
; Patent No. 6251389
;  GENERAL INFORMATION:
;     APPLICANT:  Magna, Holly
;     APPLICANT:  Schaffer, Paul
;     APPLICANT:  Lawton, Michael
;     APPLICANT:  Yocum, Sue
;     APPLICANT:  Mitchell, Peter
;     APPLICANT:  Hutchinson, Nancy
;     APPLICANT:  Murry, Lynn E.
;     TITLE OF INVENTION:  HUMAN NUCLEOTIDE PYROPHOS
;     TITLE OF INVENTION:  PHOHYDROLASE-2
;     NUMBER OF SEQUENCES:  3
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;       STREET:  3174 Porter Drive
;       CITY:  Palo Alto
;       STATE:  CA
;       COUNTRY:  USA
;       ZIP:  94304
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  DOS
;       SOFTWARE:  FastSEQ for Windows Version 2.0
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/429,516
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/996,083
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Billings, Lucy J
;       REGISTRATION NUMBER:  36,749
;       REFERENCE/DOCKET NUMBER:  PF-0420 US
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  650-855-0555
;       TELEFAX:  650-845-4166
;       TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  4183 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     IMMEDIATE SOURCE:
;       LIBRARY:  SATPF1008
;       CLONE:  1388013
US-09-429-516-2


  Query Match             26.5%;  Score 26.8;  DB 4;  Length 4183;
  Best Local Similarity   57.0%;  Pred. No. 23;
  Matches   49;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy       5 ccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggagg 64
           |||| |||| || | |||  | |||  || | |||   | |||| | ||| | |     |
Db     576 CCCCGCTGGGGATGCGTGTCCCGGGCGTCCTCTGGAGGCGCAGAAGTGCGTGCGGCCTCG 635

Qy      65 ggagtcctcctggagcctggtgccct 90
           |    |    || ||||| |   |||
Db     636 GTGTCCAGGGTGCAGCCTTGACACCT 661


RESULT  15
US-08-290-665A-140/c
; Sequence 140, Application US/08290665A
; Patent No. 5882852
;  GENERAL INFORMATION:
;     APPLICANT:  BUKH, J., MILLER, R.H. AND
;     APPLICANT:  PURCELL, R.H.
;     TITLE OF INVENTION:  NUCLEOTIDE AND DEDUCED
;     TITLE OF INVENTION:  AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
;     TITLE OF INVENTION:  CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
;     TITLE OF INVENTION:  AND THE USE OF REAGENTS DERIVED FROM THESE
;     TITLE OF INVENTION:  SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
;     NUMBER OF SEQUENCES:  263
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  MORGAN & FINNEGAN
;       STREET:  345 PARK AVENUE
;       CITY:  NEW YORK
;       STATE:  NEW YORK
;       COUNTRY:  USA
;       ZIP:  10154
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  FLOPPY DISK
;       COMPUTER:  IBM PC COMPATIBLE
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;       SOFTWARE:  WORDPERFECT 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/290,665A
;       FILING DATE:  15-AUG-1994
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  RICHARD W. BORK
;       REGISTRATION NUMBER:  36,459
;       REFERENCE/DOCKET NUMBER:  2026-4116
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 758-4800
;       TELEFAX:  (212) 751-6849
;       TELEX:  421792
;   INFORMATION FOR SEQ ID NO:  140:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  573 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     ORIGINAL SOURCE:
;       ORGANISM:  homosapiens
;       INDIVIDUAL ISOLATE:  Z8
US-08-290-665A-140


  Query Match             26.3%;  Score 26.6;  DB 2;  Length 573;
  Best Local Similarity   56.2%;  Pred. No. 21;
  Matches   50;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy       5 ccccccctggtgaggagtgggctgggaatcttatgggcacccagaggggcgggggcggagg 64
           |||   | | | |   |  || ||| ||| | ||||| || ||| || || | || | ||
Db     359 CCCAAATTGCGCGACCTCCGCCGGGGATCATTTGGGCCCCAAGACGGTCGAGAGCCGCGG 300

Qy      65 ggagtcctcctggagcctggtgccctaga 93
           || | |       | ||||    || | |
Db     299 GGGGACAGGAGCCAACCTGCCCACCCACA 271


Search completed: October  3, 2002, 16:21:45
Job time: 16945 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 16:30:18 ; Search time 719.93 Seconds
(without alignments)
3927.824 Million cell updates/sec

Title: US-09-826-581-5
Perfect score: 1647
Sequence: 1 ttggtctggggctggccaca.....acaccagcctcttagtcttc 1647

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1647	100.0	1647	22	AAH43685	PRKAG3 cDNA. Homo
	2	1453	88.2	2115	22	AAD03320	Human AMPK gamma s
	3	1447	87.9	2109	22	AAD03296	Human AMPK gamma s
	4	1172.8	71.2	2022	22	AAD03321	Sus scrofa PRKAG3
	5	1140.4	69.2	1873	22	AAD03319	Pig AMPK gamma sub
	6	1134.4	68.9	1867	22	AAD03295	Pig AMPK gamma sub
	7	448.8	27.2	547	22	ABA08485	Human AMP-activate
	8	407.2	24.7	989	22	AAH43682	PRKAG3 intron 2 -
	9	393.4	23.9	1691	21	AAC98774	Human pancreatic c
	10	390.2	23.7	1576	18	AAT85927	Mammalian AMPK-gam
	11	366.6	22.3	1435	20	AAX06882	Disease associated
	12	366.6	22.3	2223	22	AAH14839	Human cDNA sequenc
	13	299.2	18.2	3261	23	ABL18857	Drosophila melanog
	14	278.8	16.9	1014	22	AAH43684	PRKAG3 intron 10 -
c	15	252	15.3	3425	22	AAK72740	Human immune/haema
c	16	252	15.3	3425	22	AAK72741	Human immune/haema
	17	222	13.5	1467	23	AAS84265	DNA encoding novel
	18	213.2	12.9	735	22	AAH07561	Human cDNA clone (
	19	188.4	11.4	602	22	AAH35203	Human colon cancer
	20	168.2	10.2	1722	22	AAH43683	PRKAG3 intron 4 -
	21	159.8	9.7	2303	23	AAS84267	DNA encoding novel
	22	127.2	7.7	39651	23	ABL18856	Drosophila melanog
	23	120.8	7.3	350	21	AAC01661	Human secreted pro
c	24	85	5.2	378	22	ABA44706	Human breast cell
c	25	85	5.2	378	22	ABA55162	Human foetal liver
c	26	85	5.2	378	22	ABA24907	Probe #3373 for ge
c	27	85	5.2	378	22	AAK03423	Human brain expres
c	28	85	5.2	378	22	AAK28874	Human bone marrow
c	29	85	5.2	378	22	AAI13465	Probe #3398 for ge
c	30	85	5.2	378	22	AAI34821	Probe #3507 used t
c	31	85	5.2	378	22	AAI03344	Probe #3335 used t
	32	84.8	5.1	151	24	ABI99496	Mouse ischaemic co
c	33	72	4.4	16525	22	AAK73303	Human immune/haema
	34	58.8	3.6	1266	21	AAA51352	Soybean sucrose no
c	35	54	3.3	92	22	ABA49850	Human breast cell
c	36	54	3.3	92	22	ABA67769	Human foetal liver
c	37	54	3.3	92	22	ABA34826	Probe #13292 for g
c	38	54	3.3	92	22	AAK16181	Human brain expres
c	39	54	3.3	92	22	AAK41922	Human bone marrow
c	40	54	3.3	92	22	AAI22692	Probe #12625 for g
c	41	54	3.3	92	22	AAI47988	Probe #16674 used
c	42	54	3.3	92	22	AAI08354	Probe #8345 used t
c	43	53.4	3.2	426	22	AAL10935	Human breast cance
	44	52.8	3.2	523	21	AAF08150	Fusarium venenatum
c	45	52.2	3.2	415	22	AAL18884	Human breast cance

ALIGNMENTS

RESULT 1
AAH43685
ID AAH43685 standard; cDNA; 1647 BP.
XX
AC AAH43685;
XX
DT 21-JAN-2002 (first entry)
XX
DE PRKAG3 cDNA.
XX
KW Human; AMP-activated protein kinase gamma 3 subunit; PRKAG3; variant;
KW metabolic disease; diabetes; obesity; substitution; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 20..1489
FT /*tag= a
FT /product= "PRKAG3"
FT variation 230
FT /*tag= b
FT /label= "C230G"
FT /note= "Causes P71A"
FT variation 559
FT /*tag= c
FT /label= "T559C"
FT /note= "Silent variation"
FT variation 1037
FT /*tag= d
FT /label= "C1037T"
FT /note= "Causes R340W"
XX

PN WO200177305-A2.
XX
PD 18-OCT-2001.
XX
PF 06-APR-2001; 2001WO-SE00765.
XX
PR 07-APR-2000; 2000US-195665P.
XX
PA (AREX-) AREXIS AB.
XX
PI Andersson L, Luthman H, Marklund S;
XX
DR WPI; 2001-657170/75.
DR P-PSDB; QQB47679.
XX
PT New variants of human AMP-activated protein kinase gamma3 subunit
PT associated with a metabolic disease e.g. diabetes or obesity and method
PT for determining a risk estimate of diseases in subject by detecting the
PT variant -
XX
PS Disclosure; Fig 5; 25pp; English.
XX
CC This sequence represents the full length cDNA encoding the human
CC AMP-activated protein kinase gamma 3 subunit (PRKAG3). Detecting
CC the presence of the PRKAG3 DNA, or a variant, is useful in determining
CC a risk estimate of a metabolic disease, such as diabetes or obesity,
CC in a subject. The variation may occur in exons 3, 4 or 10. In exon
CC 3 variation may be a substitution of a G for a C at nucleotide 320,
CC resulting in the amino acid substitution P71A; in exon 4 variation may
CC be a substitution of a T for a C at nucleotide 550; and in exon 10
CC variation may be a substitution of a T for a C at nucleotide 1037,
CC resulting in the amino acid substitution R340W. There may also be
CC nucleotide variation in intron 6. The numbering of these
CC variations is based on the full length cDNA as given, rather than on
CC position 1 of the open reading frame.
XX
SQ Sequence 1647 BP; 346 A; 502 C; 462 G; 337 T; 0 other;

Query Match 100.0%; Score 1647; DB 22; Length 1647;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1647; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 ttggtctggggctggccacatggagcccgggctggagcacgcactgcgcaggacccccttc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ttggtctggggctggccacatggagcccgggctggagcacgcactgcgcaggacccccttc 60

Qy     61 ctggagcagccttgggggttctgagcatcaagagatgagcttcctagagcaagaaaacag 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ctggagcagccttgggggttctgagcatcaagagatgagcttcctagagcaagaaaacag 120

Qy    121 cagctcatggccatcaccagctgtgaccagcagctcagaaagaatccgtgggaaacggag 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 cagctcatggccatcaccagctgtgaccagcagctcagaaagaatccgtgggaaacggag 180

Qy    181 ggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtca 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 ggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtca 240

Qy    241 gggggaaggtccccggtccaggccagctgctgagtccaccgggctggaggccacattccc 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 gggggaaggtccccggtccaggccagctgctgagtccaccgggctggaggccacattccc 300

Qy    301 caagaccacacccttggctcaagctgatcctgccggggtgggcactccaccaacagggtg 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 caagaccacacccttggctcaagctgatcctgccggggtgggcactccaccaacagggtg 360

Qy    361 ggactgcctcccctctgactgtacagcctcagctgcaggctccagcacagatgatgtgga 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ggactgcctcccctctgactgtacagcctcagctgcaggctccagcacagatgatgtgga 420

Qy    421 gctggccacggagttcccagccacagaggcctgggagtgtgagctagaaggcctgctgga 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 gctggccacggagttcccagccacagaggcctgggagtgtgagctagaaggcctgctgga 480

Qy    481 agagaggcctgccctgtgcctgtccccgcaggccccatttcccaagctgggctgggatga 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 agagaggcctgccctgtgcctgtccccgcaggccccatttcccaagctgggctgggatga 540

Qy    541 cgaactgcggaaacccggcgcccagatctacatgcgcttcatgcaggagcacacctgcta 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 cgaactgcggaaacccggcgcccagatctacatgcgcttcatgcaggagcacacctgcta 600

Qy    601 cgatgccatggcaactagctccaagctagtcatcttcgacaccatgctggagatcaagaa 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 cgatgccatggcaactagctccaagctagtcatcttcgacaccatgctggagatcaagaa 660

Qy    661 ggccttctttgctctggtggccaacggtgtgcgggcagcccctctatgggacagcaagaa 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ggccttctttgctctggtggccaacggtgtgcgggcagcccctctatgggacagcaagaa 720

Qy    721 gcagagctttgtggggatgctgaccatcactgacttcatcctggtgctgcatcgctacta 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 gcagagctttgtggggatgctgaccatcactgacttcatcctggtgctgcatcgctacta 780

Qy    781 caggtccccccctggtccagatctatgagattgaacaacataagattgagacctggaggga 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 caggtccccccctggtccagatctatgagattgaacaacataagattgagacctggaggga 840

Qy    841 gatctacctgcaaggctgcttcaagcctctggtctccatctctcctaatgatagcctgtt 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 gatctacctgcaaggctgcttcaagcctctggtctccatctctcctaatgatagcctgtt 900

Qy    901 tgaagctgtctacaccctcatcaagaaccggatccatcgcctgcctgttcttgacccggt 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 tgaagctgtctacaccctcatcaagaaccggatccatcgcctgcctgttcttgacccggt 960

Qy    961 gtcaggcaacgtactccacatcctcacacacaaacgcctgctcaagttcctgcacatctt 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 gtcaggcaacgtactccacatcctcacacacaaacgcctgctcaagttcctgcacatctt 1020

Qy   1021 tggttccctgctgccccggccctccttcctctaccgcactatccaagatttgggcatcgg 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 tggttccctgctgccccggccctccttcctctaccgcactatccaagatttgggcatcgg 1080

Qy   1081 cacattccgagacttggctgtggtgctggagacagcacccatcctgactgcactggacat 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 cacattccgagacttggctgtggtgctggagacagcacccatcctgactgcactggacat 1140

Qy   1141 ctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtggtcaggtcgtggg 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 ctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtggtcaggtcgtggg 1200

Qy   1201 cctctattcccgctttgatgtgattcacctggctgcccagcaaacctacaaccacctgga 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 cctctattcccgctttgatgtgattcacctggctgcccagcaaacctacaaccacctgga 1260

Qy   1261 catgagtgtgggagaagccctgaggcagaggacactatgtctggagggagtcctttcctg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 catgagtgtgggagaagccctgaggcagaggacactatgtctggagggagtcctttcctg 1320

Qy   1321 ccagccccacgagagcttgggggaagtgatcgacaggattgctcgggagcaggtacacag 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 ccagccccacgagagcttgggggaagtgatcgacaggattgctcgggagcaggtacacag 1380

Qy   1381 gctggtgctagtggacgagacccagcatctcttgggcgtggtctccctctccgacatcct 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 gctggtgctagtggacgagacccagcatctcttgggcgtggtctccctctccgacatcct 1440

Qy   1441 tcaggcactggtgctcagccctgctggcatcgatgccctcggggcctgagaagatctgag 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 tcaggcactggtgctcagccctgctggcatcgatgccctcggggcctgagaagatctgag 1500

Qy   1501 tcctcaatcccaagccacctgcacacctggaagccaatgaagggaactggagaactcagc 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 tcctcaatcccaagccacctgcacacctggaagccaatgaagggaactggagaactcagc 1560
```

```
Qy   1561 cttcatcttccccacccccatttgctggttcagctatgattcaggtaggctctgccctg 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 cttcatcttccccacccccatttgctggttcagctatgattcaggtaggctctgccctg 1620

Qy   1621 ggccatgacaccagcctcttagtcttc 1647
          |||||||||||||||||||||||||||
Db   1621 ggccatgacaccagcctcttagtcttc 1647


RESULT   2
AAD03320
ID   AAD03320 standard; cDNA; 2115 BP.
XX
AC   AAD03320;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Human AMPK gamma subunit muscle-specific isoform, complete PRKAG3 cDNA.
XX
KW   Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW   PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW   genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW   cystathione beta synthase; CBS; cardiant; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1395
FT                   /*tag= a
FT                   /product= "Human complete Prkag3 protein"
XX
PN   WO200120003-A2.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-EP09896.
XX
PR   10-SEP-1999;   99EP-0402236.
PR   18-MAY-2000; 2000EP-0401388.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA   (ANDE/) ANDERSSON L.
PA   (LOOF/) LOOFT C.
PA   (KALM/) KALM E.
XX
PI   Andersson L,  Looft C,  Kalm E,  Milan D,  Robic A,  Rogel-Gaillard C;
PI   Iannuccelli N,  Gellin J,  Le Roy P,  Chardon P;
XX
DR   WPI; 2001-244810/25.
DR   P-PSDB; AAE00223.
XX
PT   New variants of the gamma subunit of vertebrate adenosine
PT   monophosphate-activated kinase for diagnosis or treatment of disorders
PT   associated with energy metabolism such as diabetes, obesity, and
PT   myopathy -
XX
PS   Claim 12; Page 65-68; 71pp; English.
XX
CC   The present sequence is a cDNA encoding human adenosine monophosphate
CC   (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC   complete PRKAG3. Mutation in Prkag3 results in an altered regulation of
CC   carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is
CC   useful as therapeutic for treating carbohydrate metabolism disorders such
CC   as diabetes, obesity, and disorders associated with muscle metabolism
CC   such as myopathy and cardiovascular diseases, to modulate AMPK
CC   activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC   and its functionally altered mutants are useful for the diagnostic
CC   evaluation, genetic testing and prognosis of a metabolic disorder,
CC   preferably a carbohydrate metabolism disorder. Primers that can detect
CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
CC   from the expression of a functionally altered allele of PRKAG3.
CC   Transgenic animal and host cell transformed with PRKAG3 or a
CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC   screening compounds able to modulate AMPK activity. Nucleic acid
CC   encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
CC   of PRKAG3 and is useful in gene therapy.
XX
SQ   Sequence 2115 BP; 460 A; 622 C; 562 G; 471 T; 0 other;


  Query Match             88.2%;  Score 1453;  DB 22;  Length 2115;
  Best Local Similarity   97.3%;  Pred. No. 0;
  Matches 1501;  Conservative    0;  Mismatches   35;  Indels    7;  Gaps    2;

Qy     95 atgagcttcctagagcaagaaaacagcagctcatggccatcaccagctgtgaccagcagc 154
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 atgagcttcctagagcaagaaaacagcagctcatggccatcaccagctgtgaccagcagc 60

Qy    155 tcagaaagaatccgtgggaaacggagggccaaagccttgagatggacaaggcagaagtcg 214
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 tcagaaagaatccgtgggaaacggagggccaaagccttgagatggacaaggcagaagtcg 120

Qy    215 gtggaggaaggggagccaccaggtcaggggggaaggtccccggtccaggccagctgctgag 274
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db    121 gtggaggaaggggagccaccaggtcaggggggaaggtccccggtccaggccaactgctgag 180

Qy    275 tccaccgggctggaggccacattccccaagaccacacccttggctcaagctgatcctgcc 334
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 tccaccgggctggaggccacattccccaagaccacacccttggctcaagctgatcctgcc 240

Qy    335 ggggtgggcactccaccaacagggtgggactgcctcccctctgactgtacagcctcagct 394
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 ggggtgggcactccaccaacagggtgggactgcctcccctctgactgtacagcctcagct 300

Qy    395 gcaggctccagcacagatgatgtggagctggccacggagttcccagccacagaggcctgg 454
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 gcaggctccagcacagatgatgtggagctggccacggagttcccagccacagaggcctgg 360

Qy    455 gagtgtgagctagaaggcctgctggaagagaggcctgccctgtgcctgtccccgcaggcc 514
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 gagtgtgagctagaaggcctgctggaagagaggcctgccctgtgcctgtccccgcaggcc 420

Qy    515 ccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcccagatctacatg 574
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 ccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcccagatctacatg 480

Qy    575 cgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagctagtcatc 634
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 cgcttca*gcaggagcacacctgctacgatgccatggcaactagctccaagctagtcatc 540

Qy    635 ttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacggtgtgcgg 694
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 ttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacggtgtgcgg 600

Qy    695 gcagcccctctatgggacagcaagaagcagagctttgtggggatgctgaccatcactgac 754
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 gcagcccctctatgggacagcaagaagcagagctttgtggggatgctgaccatcactgac 660

Qy    755 ttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatgagattgaa 814
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatgagattgaa 720

Qy    815 caacataagattgagacctggagggagatctacctgcaaggctgcttcaagcctctggtc 874
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 caacataagattgagacctggagggagatctacctgcaaggctgcttcaagcctctggtc 780

Qy    875 tccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaagaaccggatc 934
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 tccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaagaaccggatc 840

Qy    935 catcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctcacacacaaa 994
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 catcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctcacacacaaa 900
```

```
Qy     995 cgcctgctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctac 1054
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 cgcctgctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctac 960

Qy    1055 cgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgctggagaca 1114
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 cgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgctggagaca 1020

Qy    1115 gcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtg 1174
           ;||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 gcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtg 1080

Qy    1175 gtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattcacctggct 1234
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 gtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattcacctggct 1140

Qy    1235 gcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggcagaggaca 1294
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 gcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggcagaggaca 1200

Qy    1295 ctatgtctggagggagtcctttcctgccagccccacgagagcttgggggaagtgatcgac 1354
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 ctatgtctggagggagtcctttcctgccagccccacgagagcttgggggaagtgatcgac 1260

Qy    1355 aggattgctcgggagcaggtacacaggctggtgctagtggacgagacccagcatctcttg 1414
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 aggattgctcgggagcaggtacacaggctggtgctagtggacgagacccagcatctcttg 1320

Qy    1415 ggcgtggtctccctctccgacatccttcaggcactggtgctcagccctgctggcatcgat 1474
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 ggcgtggtctccctctccgacatccttcaggcactggtgctcagccctgctggcatcgat 1380

Qy    1475 gccctcggggcctgagaagatctgagtcctcaatcccaagccacctgcacacctggaagc 1534
           |||||||||||||||||||||||||||||||||||||||||| ||||||| ||||||||
Db    1381 gccctcggggcctgagaagatctgagtcctcaatcccaagccaactgcaca-ctggaagc 1439

Qy    1535 caatgaagggaactggagaactcagccttcatcttcccccaccccccatttgctggttcag 1594
           ||||||||| |    ||      |    |    |||||| ||||| ||||||||||||||
Db    1440 caatgaaggaattgaga------acagcttcatttccccaaccccaatttgctggttcag 1493

Qy    1595 ctatgattcaggtaggctctgccctgggccatgacaccagcct 1637
           |||||||||||     || ||| |     ||  |    ||||
Db    1494 ctatgattcaggcttcttcagccttccaaaattgcctttgcct 1536


RESULT   3
AAD03296
ID   AAD03296 standard; DNA; 2109 BP.
XX
AC   AAD03296;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Human AMPK gamma subunit muscle-specific isoform, PRKAG3 cDNA.
XX
KW   Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW   PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW   genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW   cystathione beta synthase; CBS; cardiant; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1..471
FT                   /*tag= a
FT   CDS             472..1389
FT                   /*tag= b
FT                   /product= "Human Prkag3 protein"
FT   3'UTR           1390..2109
FT                   /*tag= c
XX
PN   WO200120003-A2.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-EP09896.
XX
PR   10-SEP-1999;   99EP-0402236.
PR   18-MAY-2000; 2000EP-0401388.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA   (ANDE/) ANDERSSON L.
PA   (LOOF/) LOOFT C.
PA   (KALM/) KALM E.
XX
PI   Andersson L,  Looft C,  Kalm E,  Milan D,  Robic A,  Rogel-Gaillard C;
PI   Iannuccelli N,  Gellin J,  Le Roy P,  Chardon P;
XX
DR   WPI; 2001-244810/25.
DR   P-PSDB; AAE00221.
XX
PT   New variants of the gamma subunit of vertebrate adenosine
PT   monophosphate-activated kinase for diagnosis or treatment of disorders
PT   associated with energy metabolism such as diabetes, obesity, and
PT   myopathy -
XX
PS   Claim 12; Fig 2; 71pp; English.
XX
CC   The present sequence is a cDNA encoding human adenosine monophosphate
CC   (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC   PRKAG3. Mutation in Prkag3 results in an altered regulation of
CC   carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is
CC   useful as therapeutic for treating carbohydrate metabolism disorders such
CC   as diabetes, obesity, and disorders associated with muscle metabolism
CC   such as myopathy and cardiovascular diseases, to modulate AMPK
CC   activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC   and its functionally altered mutants are useful for the diagnostic
CC   evaluation, genetic testing and prognosis of a metabolic disorder,
CC   preferably a carbohydrate metabolism disorder. Primers that can detect
CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
CC   from the expression of a functionally altered allele of PRKAG3.
CC   Transgenic animal and host cell transformed with PRKAG3 or a
CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC   screening compounds able to modulate AMPK activity. Nucleic acid
CC   encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
CC   of PRKAG3 and is useful in gene therapy.
XX
SQ   Sequence 2109 BP; 458 A; 621 C; 560 G; 470 T; 0 other;


  Query Match             87.9%;  Score 1447;  DB 22;  Length 2109;
  Best Local Similarity   97.3%;  Pred. No. 0;
  Matches 1495;  Conservative    0;  Mismatches   35;  Indels    7;  Gaps    2;

Qy     101 ttcctagagcaagaaaacagcagctcatggccatcaccagctgtgaccagcagctcagaa 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ttcctagagcaagaaaacagcagctcatggccatcaccagctgtgaccagcagctcagaa 60

Qy     161 agaatccgtgggaaacggagggccaaagccttgagatggacaaggcagaagtcggtggag 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 agaatccgtgggaaacggagggccaaagccttgagatggacaaggcagaagtcggtggag 120

Qy     221 gaaggggagccaccaggtcaggggggaaggtccccggtccaggccagctgctgagtccacc 280
           |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db     121 gaaggggagccaccaggtcaggggggaaggtccccggtccaggccaactgctgagtccacc 180

Qy     281 gggctggaggccacattccccaagaccacacccttggctcaagctgatcctgccggggtg 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 gggctggaggccacattccccaagaccacacccttggctcaagctgatcctgccggggtg 240

Qy     341 ggcactccaccaacagggtgggactgcctcccctctgactgtacagcctcagctgcaggc 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 ggcactccaccaacagggtgggactgcctcccctctgactgtacagcctcagctgcaggc 300
```

```
Qy    401 tccagcacagatgatgtggagctggccacggagttcccagccacagaggcctgggagtgt 460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 tccagcacagatgatgtggagctggccacggagttcccagccacagaggcctgggagtgt 360

Qy    461 gagctagaaggcctgctggaagagaggcctgccctgtgcctgtccccgcaggccccattt 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 gagctagaaggcctgctggaagagaggcctgccctgtgcctgtccccgcaggccccattt 420

Qy    521 cccaagctgggctgggatgacgaactgcggaaacccggcgcccagatctacatgcgcttc 580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 cccaagctgggctgggatgacgaactgcggaaacccggcgcccagatctacatgcgcttc 480

Qy    581 atgcaggagcacacctgctacgatgccatggcaactagctccaagctagtcatcttcgac 640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 atgcaggagcacacctgctacgatgccatggcaactagctccaagctagtcatcttcgac 540

Qy    641 accatgctggagatcaagaaggccttctttgctctggtggccaacggtgtgcgggcagcc 700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 accatgctggagatcaagaaggccttctttgctctggtggccaacggtgtgcgggcagcc 600

Qy    701 cctctatgggacagcaagaagcagagctttgtggggatgctgaccatcactgacttcatc 760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 cctctatgggacagcaagaagcagagctttgtggggatgctgaccatcactgacttcatc 660

Qy    761 ctggtgctgcatcgctactacaggtccccctggtccagatctatgagattgaacaacat 820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 ctggtgctgcatcgctactacaggtccccctggtccagatctatgagattgaacaacat 720

Qy    821 aagattgagacctggagggagatctacctgcaaggctgcttcaagcctctggtctccatc 880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 aagattgagacctggagggagatctacctgcaaggctgcttcaagcctctggtctccatc 780

Qy    881 tctcctaatgatagcctgtttgaagctgtctacaccctcatcaagaaccggatccatcgc 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 tctcctaatgatagcctgtttgaagctgtctacaccctcatcaagaaccggatccatcgc 840

Qy    941 ctgcctgttcttgacccggtgtcaggcaacgtactccacatcctcacacacaaacgcctg 1000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 ctgcctgttcttgacccggtgtcaggcaacgtactccacatcctcacacacaaacgcctg 900

Qy   1001 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 1060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 960

Qy   1061 atccaagatttgggcatcggcacattccgagacttggctgtggtgctggagacagcaccc 1120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 atccaagatttgggcatcggcacattccgagacttggctgtggtgctggagacagcaccc 1020

Qy   1121 atcctgactgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaac 1180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 atcctgactgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaac 1080

Qy   1181 gaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattcacctggctgcccag 1240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 gaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattcacctggctgcccag 1140

Qy   1241 caaacctacaaccacctggacatgagtgtgggagaagccctgaggcagaggacactatgt 1300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 caaacctacaaccacctggacatgagtgtgggagaagccctgaggcagaggacactatgt 1200

Qy   1301 ctggagggagtcctttcctgccagccccacgagagcttgggggaagtgatcgacaggatt 1360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ctggagggagtcctttcctgccagccccacgagagcttgggggaagtgatcgacaggatt 1260

Qy   1361 gctcgggagcaggtacacaggctggtgctagtggacgagacccagcatctcttgggcgtg 1420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 gctcgggagcaggtacacaggctggtgctagtggacgagacccagcatctcttgggcgtg 1320

Qy   1421 gtctccctctccgacatccttcaggcactggtgctcagccctgctggcatcgatgccctc 1480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 gtctccctctccgacatccttcaggcactggtgctcagccctgctggcatcgatgccctc 1380

Qy   1481 ggggcctgagaagatctgagtcctcaatcccaagccacctgcacacctggaagccaatga 1540
          ||||||||||||||||||||||||||||||||||||||||| ||||||| ||||||||||||
Db   1381 ggggcctgagaagatctgagtcctcaatcccaagccaactgcaca-ctggaagccaatga 1439

Qy   1541 agggaactggagaactcagccttcatcttcccccacccccatttgctggttcagctatga 1600
          ||| |    ||      |    |    |||||| ||||| ||||||||||||||||||||
Db   1440 aggaattgaga------acagcttcatttccccaaccccaatttgctggttcagctatga 1493

Qy   1601 ttcaggtaggctctgccctgggccatgacaccagcct 1637
          ||||||     || ||| |     ||  |    ||||
Db   1494 ttcaggcttcttcagccttccaaaattgcctttgcct 1530
```

RESULT 4
AAD03321
ID AAD03321 standard; DNA; 2022 BP.
XX
AC AAD03321;
XX
DT 13-JUN-2001 (first entry)
XX
DE Sus scrofa PRKAG3 splice variant DNA.
XX
KW Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW cystathione beta synthase; CBS; cardiant; gene therapy; ds.
XX
OS Sus scrofa.
XX
FH Key Location/Qualifiers
FT CDS 1..1545
FT /*tag= a
FT /product= "Sus scrofa Prkag3 splice variant"
XX
PN WO200120003-A2.
XX
PD 22-MAR-2001.
XX
PF 11-SEP-2000; 2000WO-EP09896.
XX
PR 10-SEP-1999; 99EP-0402236.
PR 18-MAY-2000; 2000EP-0401388.
XX
PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
PA (ANDE/) ANDERSSON L.
PA (LOOF/) LOOFT C.
PA (KALM/) KALM E.
XX
PI Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;
PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;
XX
DR WPI; 2001-244810/25.
DR P-PSDB; AAE00224.
XX
PT New variants of the gamma subunit of vertebrate adenosine
PT monophosphate-activated kinase for diagnosis or treatment of disorders
PT associated with energy metabolism such as diabetes, obesity, and
PT myopathy -
XX
PS Claim 12; Page 69; 71pp; English.
XX
CC The present sequence is pig adenosine monophosphate (AMP)-activated
CC kinase (AMPK) gamma subunit muscle-specific isoform, PRKAG3 splice
CC variant DNA. Prkag3 gene is located in the RN locus of chromosome 15.
CC Mutation in Prkag3 results in an altered regulation of carbohydrate
CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as
CC therapeutic for treating carbohydrate metabolism disorders such as
CC diabetes, obesity, and disorders associated with muscle metabolism
CC such as myopathy and cardiovascular diseases, to modulate AMPK
CC activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC and its functionally altered mutants are useful for the diagnostic
CC evaluation, genetic testing and prognosis of a metabolic disorder,

```
CC   preferably a carbohydrate metabolism disorder. Primers that can detect
CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
CC   from the expression of a functionally altered allele of PRKAG3.
CC   Transgenic animal and host cell transformed with PRKAG3 or a
CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC   screening compounds able to modulate AMPK activity. Nucleic acid
CC   encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
CC   of PRKAG3 and is useful in gene therapy.
XX
SQ   Sequence 2022 BP; 412 A; 623 C; 593 G; 394 T; 0 other;


  Query Match             71.2%;  Score 1172.8;  DB 22;  Length 2022;
  Best Local Similarity   85.7%;  Pred. No. 4e-296;
  Matches 1351;  Conservative    0;  Mismatches  217;  Indels    8;  Gaps    4;

Qy      53 acccctcctggagcagccttgggggttctgagcatcaagagatgagcttcctagagcaa 112
           || || ||||||||||||||| |||||  || ||||||||||||||||||||||||||||
Db     109 actccctcctggagcagcctcggggacctaagcatcaagagatgagcttcctagagcaa 168

Qy     113 gaaaacagcagctcatggccatcaccagctgtgaccagcagctcagaaagaatccgtggg 172
           | | | ||| | |||||||||||| | |||||| |||| |||||||||||||| || ||||
Db     169 ggagagagccgttcatggccatcccgagctgtaaccaccagctcagaaagaagccatggg 228

Qy     173 aaacggagggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagcca 232
             | | | ||  ||| ||||  |||||||||||||||  ||   || ||||||||| |||
Db     229 gaccaggggaacaaggcctctagatggacaaggcaggaggatgtagaggaaggggggcct 288

Qy     233 ccaggtcaggggggaaggtccccggtccaggccagctgctgagtccaccgggctggaggcc 292
           || || | | |||||||||||| |||||||||| |||||||||||||||| |||||||
Db     289 ccgggcccgagggaaggtccccagtccaggccagttgctgagtccaccgggcaggaggcc 348

Qy     293 acattccccaagaccacacccttggctcaagctgatcc---tgccggggtgggcactcca 349
           |||||||||||| |||||||||||| ||||| | |||    |||| ||||| ||  ||
Db     349 acattccccaaggccacacccttggcccaagccgctcccttggccgaggtggacaacccc 408

Qy     350 ccaacagggtgggactgcctcccctctgactgtacagcctcagctgcaggctccagcaca 409
           ||||||| | |||||  |||||||||||||||| |||||||||  | | ||||| ||||
Db     409 ccaacagagcgggacatcctcccctctgactgtgcagcctcagcctccgactccaacaca 468

Qy     410 gatgatgtggagctggccacggagttcccagccacagaggcctgggagtgtgagctagaa 469
           ||  || |||| |||| ||  |||||| ||||| | | ||| | || |  ||||||
Db     469 gaccatctggatctgggcatagagttctcagcctcggcggcgtcgggggatgagct---t 525

Qy     470 ggcctgctggaagagaggcctgccctgtgcctgtccccgcaggccccatttcccaagctg 529
           || ||| ||||||||| ||| |||| |||||  |||||  |||  |  || |||| ||||
Db     526 gggctggtggaagagaagccagccccgtgcccatccccagaggtgctgttacccaggctg 585

Qy     530 ggctgggatgacgaactgcggaaaccccggcgcccagatctacatgcgcttcatgcaggag 589
           |||||||||| || |||| ||| || || |||||| ||||||||| ||||||||||||||
Db     586 ggctgggatgatgagctgcagaagccgggggcccaggtctacatgcacttcatgcaggag 645

Qy     590 cacacctgctacgatgccatggcaactagctccaagctagtcatcttcgacaccatgctg 649
           |||||||||||||||||||||| || |||||||| || ||||||||||||||||||||
Db     646 cacacctgctacgatgccatggcgaccagctccaaactggtcatcttcgacaccatgctg 705

Qy     650 gagatcaagaaggccttctttgctctggtggccaacggtgtgcgggcagcccctctatgg 709
           |||||||||||||||||||||| |||||||||||||| || || || || ||| | |||
Db     706 gagatcaagaaggccttctttgccctggtggccaacggcgtccgagcggcacctttgtgg 765

Qy     710 gacagcaagaagcagagctttgtggggatgctgaccatcactgacttcatcctggtgctg 769
           |||||||||||||||||| |||||||||||||||||| ||||||||| ||||||||
Db     766 gacagcaagaagcagagcttcgtggggatgctgaccatcacagacttcatcttggtgctg 825

Qy     770 catcgctactacaggtccccccctggtccagatctatgagattgaacaacataagattgag 829
           || ||||| |||||||||||||||||||||||||||| ||||||||| ||||||||||||||
Db     826 caccgctattacaggtccccccctggtccagatctacgagattgaagaacataagattgag 885

Qy     830 acctggagggagatctacctgcaaggctgcttcaagcctctggtctccatctctcctaat 889
           |||||||||||||||||| |||||||||||||||||||||||||||||||||||||| |||
Db     886 acctggagggagatctaccttcaaggctgcttcaagcctctggtctccatctctcccaat 945

Qy     890 gatagcctgtttgaagctgtctacaccctcatcaagaaccggatccatcgcctgcctgtt 949
           || |||||||| |||||||||||| |||||||||||||||||||||| |||||||| ||
Db     946 gacagcctgttcgaagctgtctacgccctcatcaagaaccggatccaccgcctgccggtc 1005

Qy     950 cttgacccggtgtcaggcaacgtactccacatcctcacacacaaacgcctgctcaagttc 1009
           || ||||| || || ||    || ||||||||||||||||| || || || |||||||||
Db    1006 ctggaccctgtctccggggctgtgctccacatcctcacacataagcggcttctcaagttc 1065

Qy    1010 ctgcacatctttggttccctgctgccccggccctccttcctctaccgcactatccaagat 1069
           |||||||||||||  |||||||||||||||||||||||||||||||||| |||||||||
Db    1066 ctgcacatctttggcaccctgctgccccggccctccttcctctaccgcaccatccaagat 1125

Qy    1070 ttgggcatcggcacattccgagacttggctgtggtgctggagacagcacccatcctgact 1129
           |||||||||||||||||||||||||||||| |||||||||| || || ||||||||||
Db    1126 ttgggcatcggcacattccgagacttggccgtggtgctggaaacggcgcccatcctgacc 1185

Qy    1130 gcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtggt 1189
           |||||||||||||| |||||||||||||||||| |||||||||||||||||  |||
Db    1186 gcactggacatcttcgtggaccggcgtgtgtctgcgctgcctgtggtcaacgaaactgga 1245

Qy    1190 caggtcgtgggcctctattcccgctttgatgtgattcacctggctgcccagcaaacctac 1249
           ||||| |||||||||| || |||||||||||||| |||||||||||||||| ||||| |||
Db    1246 caggtagtgggcctctactctcgctttgatgtgatccacctggctgcccaacaaacatac 1305

Qy    1250 aaccacctggacatgagtgtgggagaagccctgaggcagaggacactatgtctggaggga 1309
           |||||||||||||||| |||||||||||||||||||||| ||||||| |||||||| ||
Db    1306 aaccacctggacatgaatgtgggagaagccctgaggcagcggacactgtgtctggaaggc 1365

Qy    1310 gtcctttcctgccagccccacgagagcttgggggaagtgatcgacaggattgctcgggag 1369
           |||||||||||||||||||||||||| |||||||||||| || ||| ||||||  |||||
Db    1366 gtcctttcctgccagccccacgagaccttgggggaagtcattgaccggattgtccgggaa 1425

Qy    1370 caggtacacaggctggtgctagtggacgagacccagcatctcttgggcgtggtctccctc 1429
           ||||| ||| | |||||||| ||||| |||||||||| ||  |||||||||| ||||||
Db    1426 caggtgcaccgcctggtgctcgtggatgagacccagcaccttctgggcgtggtgtccctc 1485

Qy    1430 tccgacatccttcaggcactggtgctcagccctgctggcatcgatgccctcggggcctga 1489
           || |||||||||||||| |||||||||||||||||||| || ||||||||||||||||||
Db    1486 tctgacatccttcaggctctggtgctcagccctgctggaattgatgccctcggggcctga 1545

Qy    1490 gaagatctgagtcctcaatcccaagccacct-gcacacctggaagccaatgaagggaact 1548
           |||  |  ||  | |   || || |||||||  |||||||||||||||| |||||||| |
Db    1546 gaaccttggaacctttgctctcaggccacctggcacacctggaagccagtgaagggagc- 1604

Qy    1549 ggagaactcagccttcatcttcccccaccccccatttgctggttcagctatgattcaggta 1608
             | | |||||||  |||  | ||| || ||||| |||||||||   ||| |  ||||||||
Db    1605 cgtggactcagctctcacttcccctcagccccacttgctggtctggctcttgttcaggta 1664

Qy    1609 ggctctgccctgggcc 1624
           ||||| |||| |||||
Db    1665 ggctccgcccggggcc 1680


RESULT   5
AAD03319
ID   AAD03319 standard; cDNA; 1873 BP.
XX
AC   AAD03319;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Pig AMPK gamma subunit muscle-specific isoform, complete PRKAG3 cDNA.
XX
KW   Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW   PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW   genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW   cystathione beta synthase; CBS; cardiant; gene therapy; RN locus;
KW   chromosome 15; ss.
XX
OS   Sus scrofa.
XX
FH   Key             Location/Qualifiers
```

```
FT   CDS             1..1395
FT                   /*tag= a
FT                   /product= "Sus scrofa complete Prkag3 protein"
XX
PN   WO200120003-A2.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-EP09896.
XX
PR   10-SEP-1999;   99EP-0402236.
PR   18-MAY-2000; 2000EP-0401388.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA   (ANDE/) ANDERSSON L.
PA   (LOOF/) LOOFT C.
PA   (KALM/) KALM E.
XX
PI   Andersson L,  Looft C,  Kalm E,  Milan D,  Robic A,  Rogel-Gaillard C;
PI   Iannuccelli N,  Gellin J,  Le Roy P,  Chardon P;
XX
DR   WPI; 2001-244810/25.
DR   P-PSDB; AAE00222.
XX
PT   New variants of the gamma subunit of vertebrate adenosine
PT   monophosphate-activated kinase for diagnosis or treatment of disorders
PT   associated with energy metabolism such as diabetes, obesity, and
PT   myopathy -
XX
PS   Claim 12; Page 62-64; 71pp; English.
XX
CC   The present sequence is a cDNA encoding pig adenosine monophosphate
CC   (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC   complete PRKAG3. Prkag3 gene is located in the RN locus of chromosome
CC   15. Mutation in Prkag3 results in an altered regulation of carbohydrate
CC   metabolism, particularly in skeletal muscle. PRKAG3 is useful as
CC   therapeutic for treating carbohydrate metabolism disorders such as
CC   diabetes, obesity, and disorders associated with muscle metabolism
CC   such as myopathy and cardiovascular diseases, to modulate AMPK
CC   activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC   and its functionally altered mutants are useful for the diagnostic
CC   evaluation, genetic testing and prognosis of a metabolic disorder,
CC   preferably a carbohydrate metabolism disorder. Primers that can detect
CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
CC   from the expression of a functionally altered allele of PRKAG3.
CC   Transgenic animal and host cell transformed with PRKAG3 or a
CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC   screening compounds able to modulate AMPK activity. Nucleic acid
CC   encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
CC   of PRKAG3 and is useful in gene therapy.
XX
SQ   Sequence 1873 BP; 382 A; 580 C; 535 G; 376 T; 0 other;

  Query Match           69.2%;  Score 1140.4;  DB 22;  Length 1873;
  Best Local Similarity  85.7%;  Pred. No. 1.1e-287;
  Matches 1315;  Conservative    0;  Mismatches  211;  Indels    8;  Gaps    4;

Qy      95 atgagcttcctagagcaagaaaacagcagctcatggccatcaccagctgtgaccagcagc 154
           ||||||||||||||||||| | | ||| | |||||||||||| | |||||| |||| ||||
Db       1 atgagcttcctagagcaaggagagagccgttcatggccatcccgagctgtaaccaccagc 60

Qy     155 tcagaaagaatccgtgggaaacggagggccaaagccttgagatggacaaggcagaagtcg 214
           |||||||||| || |||| | | | ||  ||| ||||  ||||||||||||||| ||
Db      61 tcagaaagaagccatggggaccaggggaacaaggcctctagatggacaaggcaggaggat 120

Qy     215 gtggaggaaggggagccaccaggtcaggggaaggtccccggtccaggccagctgctgag 274
           || |||||||||| ||| || || | | |||||||||||| |||||||||| |||||||
Db     121 gtagaggaaggggggcctccgggcccgagggaaggtccccagtccaggccagttgctgag 180

Qy     275 tccaccgggctggaggccacattccccaagaccacacccttggctcaagctgatcc---t 331
           |||||||||| |||||||||||||||||||| |||||||||||| ||||| | |||
Db     181 tccaccgggcaggaggccacattccccaaggccacacccttggcccaagccgctcccttg 240

Qy     332 gccggggtgggcactccaccaacagggtgggactgcctcccctctgactgtacagcctca 391
           |||| ||||| ||  || ||||||| | |||||  ||||||||||||||||| ||||||||
Db     241 gccgaggtggacaacccccccaacagagcgggacatcctcccctctgactgtgcagcctca 300

Qy     392 gctgcaggctccagcacagatgatgtggagctggccacggagttcccagccacagaggcc 451
           ||  | | ||||| ||||||  || |||| |||| ||  |||||| ||||| | | |||
Db     301 gcctccgactccaacacagaccatctggatctgggcatagagttctcagcctcggcggcg 360

Qy     452 tgggagtgtgagctagaaggcctgctggaagagaggcctgccctgtgcctgtccccgcag 511
           | || |  ||||||    || ||| |||||||| ||| |||| ||||  |||||  ||
Db     361 tcgggggatgagct---tgggctggtggaagagaagccagccccgtgcccatccccagag 417

Qy     512 gccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcccagatctac 571
           |  |  || |||| ||||||||||||||| || |||| ||| || || |||||| |||||
Db     418 gtgctgttacccaggctgggctgggatgatgagctgcagaagccgggggcccaggtctac 477

Qy     572 atgcgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagctagtc 631
           |||| ||||||||||||||||||||||||||||||||||| || |||||||| || |||
Db     478 atgcacttcatgcaggagcacacctgctacgatgccatggcgaccagctccaaactggtc 537

Qy     632 atcttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacggtgtg 691
           |||||||||||||||||||||||||||||||||||||||||||| ||||||||||||| ||
Db     538 atcttcgacaccatgctggagatcaagaaggccttctttgccctggtggccaacggcgtc 597

Qy     692 cgggcagcccctctatgggacagcaagaagcagagctttgtggggatgctgaccatcact 751
           || || || ||| | |||||||||||||||||||||| ||||||||||||||||||||
Db     598 cgagcggcaccttttgtgggacagcaagaagcagagcttcgtggggatgctgaccatcaca 657

Qy     752 gacttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatgagatt 811
           ||||||||| |||||||||| ||||| |||||||||||||||||||||||||| ||||||
Db     658 gacttcatcttggtgctgcaccgctattacaggtccccccctggtccagatctacgagatt 717

Qy     812 gaacaacataagattgagacctggagggagatctacctgcaaggctgcttcaagcctctg 871
           ||| ||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db     718 gaagaacataagattgagacctggagggagatctaccttcaaggctgcttcaagcctctg 777

Qy     872 gtctccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaagaaccgg 931
           |||||||||||||| ||||| |||||||| |||||||||||| ||||||||||||||||
Db     778 gtctccatctctcccaatgacagcctgttcgaagctgtctacgccctcatcaagaaccgg 837

Qy     932 atccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctcacacac 991
           ||||| |||||||| || || ||||| || || ||    || ||||||||||||||||
Db     838 atccaccgcctgccggtcctggaccctgtctccgggggctgtgctccacatcctcacacat 897

Qy     992 aaacgcctgctcaagttcctgcacatctttggttccctgctgccccggccctccttcctc 1051
           || || || ||||||||||||||||||||||||  ||||||||||||||||||||||||||
Db     898 aagcggcttctcaagttcctgcacatctttggcaccctgctgccccggccctccttcctc 957

Qy    1052 taccgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgctggag 1111
           |||||||| ||||||||||||||||||||||||||||||||||||| ||||||||||
Db     958 taccgcaccatccaagatttgggcatcggcacattccgagacttggccgtggtgctggaa 1017

Qy    1112 acagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcactgcct 1171
           || || |||||||||||| |||||||||||||| ||||||||||||||||||||| ||||||
Db    1018 acggcgcccatcctgaccgcactggacatcttcgtggaccggcgtgtgtctgcgctgcct 1077

Qy    1172 gtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattcacctg 1231
           |||||||||||  ||| ||||| |||||||||||| || ||||||||||||||| ||||||
Db    1078 gtggtcaacgaaactggacaggtagtgggcctctactctcgctttgatgtgatccacctg 1137

Qy    1232 gctgcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggcagagg 1291
           |||||||| ||||| |||||||||||||||||| ||||||||||||||||||||||| ||
Db    1138 gctgcccaacaaacatacaaccacctggacatgaatgtgggagaagccctgaggcagcgg 1197

Qy    1292 acactatgtctggagggagtcctttcctgccagccccacgagagcttgggggaagtgatc 1351
           ||||| |||||||| || ||||||||||||||||||||||||||| |||||||||||| ||
Db    1198 acactgtgtctggaaggcgtcctttcctgccagccccacgagaccttgggggaagtcatt 1257

Qy    1352 gacaggattgctcgggagcaggtacacaggctggtgctagtggacgagacccagcatctc 1411
           ||| ||||||  ||||| ||||| ||| | |||||||| ||||| ||||||||||| ||
Db    1258 gacaggattgctcgggagcaggtacacaggctggtgctagtggacgagacccagcatctc 1317
```

```
Db   1258 gaccggattgtccgggaacaggtgcaccgcctggtgctcgtggatgagacccagcacctt 1317

Qy   1412 ttgggcgtggtctccctctccgacatccttcaggcactggtgctcagccctgctggcatc 1471
          |||||||||| |||||||| |||||||||||||| ||||||||||||||||||||| ||
Db   1318 ctgggcgtggtgtccctctctgacatccttcaggctctggtgctcagccctgctggaatt 1377

Qy   1472 gatgccctcggggcctgagaagatctgagtcctcaatcccaagccacct-gcacacctgg 1530
          |||||||||||||||||||||  |  ||  | |   || || ||||||| ||||||||||
Db   1378 gatgccctcggggcctgagaaccttggaacctttgctctcaggccacctggcacacctgg 1437

Qy   1531 aagccaatgaagggaactggagaactcagccttcatcttcccccaccccccatttgctggt 1590
          |||||| |||||||| |  | | |||||||  |||  | ||| || ||||| ||||||||
Db   1438 aagccagtgaagggagc-cgtggactcagctctcacttcccctcagccccacttgctggt 1496

Qy   1591 tcagctatgattcaggtaggctctgccctgggcc 1624
             ||| |  |||||||||||||| |||| |||||
Db   1497 ctggctcttgttcaggtaggctccgcccggggcc 1530


RESULT   6
AAD03295
ID   AAD03295 standard; cDNA; 1867 BP.
XX
AC   AAD03295;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Pig AMPK gamma subunit muscle-specific isoform, PRKAG3 cDNA.
XX
KW   Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW   PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW   genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW   cystathione beta synthase; CBS; cardiant; gene therapy; RN locus;
KW   chromosome 15; ss.
XX
OS   Sus scrofa.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1..471
FT                   /*tag= a
FT   CDS             472..1389
FT                   /*tag= b
FT                   /product= "Sus scrofa PRKAG3 protein"
FT   3'UTR           1390..1867
FT                   /*tag= c
XX
PN   WO200120003-A2.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-EP09896.
XX
PR   10-SEP-1999;   99EP-0402236.
PR   18-MAY-2000; 2000EP-0401388.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA   (ANDE/) ANDERSSON L.
PA   (LOOF/) LOOFT C.
PA   (KALM/) KALM E.
XX
PI   Andersson L,  Looft C,  Kalm E,  Milan D,  Robic A,  Rogel-Gaillard C;
PI   Iannuccelli N,  Gellin J,  Le Roy P,  Chardon P;
XX
DR   WPI; 2001-244810/25.
DR   P-PSDB; AAE00220.
XX
PT   New variants of the gamma subunit of vertebrate adenosine
PT   monophosphate-activated kinase for diagnosis or treatment of disorders
PT   associated with energy metabolism such as diabetes, obesity, and
PT   myopathy  -
XX
PS   Claim 12; Fig 2; 71pp; English.
XX
CC   The present sequence is a cDNA encoding pig adenosine monophosphate
CC   (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC   PRKAG3. Prkag3 gene is located in the RN locus of chromosome 15.
CC   Mutation in Prkag3 results in an altered regulation of carbohydrate
CC   metabolism, particularly in skeletal muscle. PRKAG3 is useful as
CC   therapeutic for treating carbohydrate metabolism disorders such as
CC   diabetes, obesity, and disorders associated with muscle metabolism
CC   such as myopathy and cardiovascular diseases, to modulate AMPK
CC   activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC   and its functionally altered mutants are useful for the diagnostic
CC   evaluation, genetic testing and prognosis of a metabolic disorder,
CC   preferably a carbohydrate metabolism disorder. Primers that can detect
CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
CC   from the expression of a functionally altered allele of PRKAG3.
CC   Transgenic animal and host cell transformed with PRKAG3 or a
CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC   screening compounds able to modulate AMPK activity. Nucleic acid
CC   encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
CC   of PRKAG3 and is useful in gene therapy.
XX
SQ   Sequence 1867 BP; 380 A; 583 C; 529 G; 375 T; 0 other;


  Query Match           68.9%;  Score 1134.4;  DB 22;  Length 1867;
  Best Local Similarity  85.7%;  Pred. No. 4.1e-286;
  Matches 1309;  Conservative    0;  Mismatches  211;  Indels    8;  Gaps    4;

Qy    101 ttcctagagcaagaaaacagcagctcatggccatcaccagctgtgaccagcagctcagaa 160
          |||||||||||||| | | ||| | |||||||||||| | |||||| |||| ||||||||||
Db      1 ttcctagagcaaggagagagccgttcatggccatcccgagctgtaaccaccagctcagaa 60

Qy    161 agaatccgtgggaaacggagggccaaagccttgagatggacaaggcagaagtcggtggag 220
          |||| || |||| | | | ||  ||| ||||  ||||||||||||||| ||   || |||
Db     61 agaagccatggggaccaggggaacaaggcctctagatggacaaggcaggaggatgtagag 120

Qy    221 gaaggggagccaccaggtcaggggaaggtccccggtccaggccagctgctgagtccacc 280
          ||||||| ||| || || | | |||||||||||| |||||||||| | ||||||||||||
Db    121 gaaggggggcctccgggcccgagggaaggtccccagtccaggccagttgctgagtccacc 180

Qy    281 gggctggaggccacattccccaagaccacacccttggctcaagctgatcc---tgccggg 337
          |||| |||||||||||||||||||| |||||||||||| ||||| | |||    |||| |
Db    181 gggcaggaggccacattccccaaggccacacccttggcccaagccgctcccttggccgag 240

Qy    338 gtgggcactccaccaacagggtgggactgcctcccctctgactgtacagcctcagctgca 397
          |||| ||  || ||||||| | |||||  |||||||||||||||| ||||||||||  |
Db    241 gtggacaacccccccaacagagcgggacatcctcccctctgactgtgcagcctcagcctcc 300

Qy    398 ggctccagcacagatgatgtggagctggccacggagttcccagccacagaggcctgggag 457
          | ||||| ||||||  || |||| |||| ||  |||||| ||||| | | ||| | || |
Db    301 gactccaacacagaccatctggatctgggcatagagttctcagcctcggcggcgtcgggg 360

Qy    458 tgtgagctagaaggcctgctggaagagaggcctgccctgtgcctgtccccgcaggcccca 517
            ||||||    || ||| |||||||||| ||| |||| |||||  |||||  |||  |
Db    361 gatgagct---tgggctggtggaagagaagccagccccgtgcccatccccagaggtgctg 417

Qy    518 tttcccaagctgggctgggatgacgaactgcggaaacccggcgcccagatctacatgcgc 577
          || |||| |||||||||||||||| || |||| ||| || || |||||| |||||||| |
Db    418 ttacccaggctgggctgggatgatgagctgcagaagccgggggcccaggtctacatgcac 477

Qy    578 ttcatgcaggagcacacctgctacgatgccatggcaactagctccaagctagtcatcttc 637
          |||||||||||||||||||||||||||||||||||| || |||||||| || |||||||||
Db    478 ttcatgcaggagcacacctgctacgatgccatggcgaccagctccaaactggtcatcttc 537

Qy    638 gacaccatgctggagatcaagaaggccttctttgctctggtggccaacggtgtgcgggca 697
          |||||||||||||||||||||||||||||||||||| ||||||||||||| || || ||
Db    538 gacaccatgctggagatcaagaaggccttctttgccctggtggccaacggcgtccgagcg 597

Qy    698 gcccctctatgggacagcaagaagcagagctttgtggggatgctgaccatcactgacttc 757
          || ||| | |||||||||||||||||||||||| |||||||||||||||||||| ||||||
Db    598 gcacctttgtgggacagcaagaagcagagcttcgtggggatgctgaccatcacagacttc 657
```

```
Qy    758 atcctggtgctgcatcgctactacaggtccccccctggtccagatctatgagattgaacaa 817
          ||| |||||||||| ||||| ||||||||||||||||||||||||| ||||||||| ||
Db    658 atcttggtgctgcaccgctattacaggtccccccctggtccagatctacgagattgaagaa 717

Qy    818 cataagattgagacctggagggagatctacctgcaaggctgcttcaagcctctggtctcc 877
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    718 cataagattgagacctggagggagatctaccttcaaggctgcttcaagcctctggtctcc 777

Qy    878 atctctcctaatgatagcctgtttgaagctgtctacaccctcatcaagaaccggatccat 937
          |||||||| ||||| |||||||| |||||||||||| ||||||||||||||||||||||
Db    778 atctctcccaatgacagcctgttcgaagctgtctacgccctcatcaagaaccggatccac 837

Qy    938 cgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctcacacacaaacgc 997
          |||||||| || || ||||| || || ||    || ||||||||||||||| || ||
Db    838 cgcctgccggtcctggaccctgtctccggggctgtgctccacatcctcacacataagcgg 897

Qy    998 ctgctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgc 1057
          || |||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db    898 cttctcaagttcctgcacatctttggcaccctgctgccccggccctccttcctctaccgc 957

Qy   1058 actatccaagatttgggcatcggcacattccgagacttggctgtggtgctggagacagca 1117
          || ||||||||||||||||||||||||||||||||||||||||| |||||||||| || ||
Db    958 accatccaagatttgggcatcggcacattccgagacttggccgtggtgctggaaacggcg 1017

Qy   1118 cccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcactgcctgtggtc 1177
          |||||||||| |||||||||||| |||||||||||||||||| ||||||||||||
Db   1018 cccatcctgaccgcactggacatcttcgtggaccggcgtgtgtctgcgctgcctgtggtc 1077

Qy   1178 aacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattcacctggctgcc 1237
          ||||||  ||| ||||| |||||||||| || |||||||||||||| ||||||||||||
Db   1078 aacgaaactggacaggtagtgggcctctactctcgctttgatgtgatccacctggctgcc 1137

Qy   1238 cagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggcagaggacacta 1297
          || ||||| |||||||||||||||||| ||||||||||||||||||||| |||||||
Db   1138 caacaaacatacaaccacctggacatgaatgtgggagaagccctgaggcagcggacactg 1197

Qy   1298 tgtctggagggagtcctttcctgccagccccacgagagcttgggggaagtgatcgacagg 1357
          |||||||| || |||||||||||||||||||||||| |||||||||||| || ||| ||
Db   1198 tgtctggaaggcgtcctttcctgccagccccacgagaccttgggggaagtcattgaccgg 1257

Qy   1358 attgctcgggagcaggtacacaggctggtgctagtggacgagacccagcatctcttgggc 1417
          ||||  ||||| ||||| ||| | |||||||| ||||| |||||||||| ||  |||||
Db   1258 attgtccgggaacaggtgcaccgcctggtgctcgtggatgagacccagcaccttctgggc 1317

Qy   1418 gtggtctccctctccgacatccttcaggcactggtgctcagccctgctggcatcgatgcc 1477
          ||||| |||||||| |||||||||||||| |||||||||||||||||||| || ||||||
Db   1318 gtggtgtccctctctgacatccttcaggctctggtgctcagccctgctggaattgatgcc 1377

Qy   1478 ctcggggcctgagaagatctgagtcctcaatcccaagccacct-gcacacctggaagcca 1536
          ||||||||||||||  |  ||  | |   || || ||||||| ||||||||||||||||
Db   1378 ctcggggcctgagaaccttggaacctttgctctcaggccacctggcacacctggaagcca 1437

Qy   1537 atgaagggaactggagaactcagccttcatcttcccccaccccccatttgctggttcagct 1596
           |||||||| |  | | ||||||| |||  | ||| || ||||| |||||||   |||
Db   1438 gtgaagggagc-cgtggactcagctctcacttcccctcagccccacttgctggtctggct 1496

Qy   1597 atgattcaggtaggctctgccctgggcc 1624
           |  |||||||||||||| |||| |||||
Db   1497 cttgttcaggtaggctccgcccgggcc 1524


RESULT   7
ABA08485
ID   ABA08485 standard; cDNA; 547 BP.
XX
AC   ABA08485;
XX
DT   11-JAN-2002  (first entry)
XX
DE   Human AMP-activated protein kinase subunit homologue cDNA, SEQ ID NO:261.
XX
KW   Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW   haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW   inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW   proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW   myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW   chronic inflammatory condition; proliferative retinopathy;
KW   atherosclerosis; coronary heart disease; arterial ischaemia;
KW   bone disorder; osteoporosis; vascular growth disorder;
KW   tissue regeneration; wound healing; infection; immune disorder;
KW   cell culture; drug screening; gene therapy; antiinflammatory;
KW   antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW   cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW   antifungal; vulnerary; antiulcer; ss.
XX
OS   Homo sapiens.
XX
PN   WO200157188-A2.
XX
PD   09-AUG-2001.
XX
PF   05-FEB-2001; 2001WO-US03800.
XX
PR   03-FEB-2000; 2000US-0496914.
PR   27-APR-2000; 2000US-0560875.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-457740/49.
DR   P-PSDB; ABB11241.
XX
PT   Human proteins and DNA encoding sequences useful for preventing,
PT   treating or ameliorating a medical condition in a mammalian subject
PT   e.g. arthritis and cancer -
XX
PS   Claim 1; Page 429; 1963pp; English.
XX
CC   Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC   sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC   invention also relates to vectors and recombinant host cells comprising a
CC   nucleotide of the invention, methods of producing the novel polypeptides,
CC   antibodies against the polypeptides, methods of detecting the nucleotides
CC   or polypeptides in a sample, and methods of identifying compounds which
CC   bind to polypeptides of the invention. Although novel, many of the
CC   polypeptides of the invention have homology to known proteins, thereby
CC   giving an insight into their probable biological activities, and hence
CC   potential therapeutic applications. The polypeptides of the invention may
CC   have various activities, including cytokine, cell proliferation or cell
CC   differentiation activities; stem cell growth factor activity;
CC   haematopoiesis regulatory activity; tissue growth activity;
CC   immunomodulatory activity; activin- or inhibin-related activities;
CC   chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC   thrombolytic activities; receptor or ligand activities; or may be
CC   involved in oncogenesis, cancer cell proliferation or metastasis.
CC   Depending on their biological activities, polypeptides and nucleotides of
CC   the invention are useful for preventing, treating or ameliorating medical
CC   conditions, e.g., by protein or gene therapy. Such conditions include
CC   cancers, haematopoietic disorders (e.g.,myeloid or lymphoid cell
CC   disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC   proliferative retinopathy, atherosclerosis, coronary heart disease,
CC   arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC   vascular growth. Polypeptides involved with tissue regeneration and
CC   repair (or nucleic acids encoding them) may be used to promote wound
CC   healing (e.g., of burns, incisions and ulcers), while those with
CC   immunomodulatory activities may be used in the treatment of viral,
CC   bacterial and fungal infections in addition to immune disorders.
CC   Polypeptides with growth factor activity may be used in cell cultures to
CC   promote cell growth. For example, such polypeptides may be used to
CC   manipulate stem cells in culture to give rise to neuroepithelial cells
CC   that can be used to augment or replace cells damaged by illness,
CC   autoimmune disease or accidental damage. The polypeptides and nucleotides
CC   may also be used in the diagnosis of the above conditions, and in drug
CC   screening techniques. The present sequence represents a cDNA encoding a
CC   novel human polypeptide of the invention.
```

```
XX
SQ   Sequence 547 BP; 112 A; 172 C; 133 G; 130 T; 0 other;


  Query Match             27.2%;  Score 448.8;  DB 22;  Length 547;
  Best Local Similarity   99.6%;  Pred. No. 2.7e-107;
  Matches  450;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy     738 tgctgaccatcactgacttcatcctggtgctgcatcgctactacaggtccccccctggtcc 797
           ||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db       1 tgctgaccatcactgacttcatcctggtgctgtatcgctactacaggtccccccctggtcc 60

Qy     798 agatctatgagattgaacaacataagattgagacctggagggagatctacctgcaaggct 857
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 agatctatgagattgaacaacataagattgagacctggagggagatctacctgcaaggct 120

Qy     858 gcttcaagcctctggtctccatctctcctaatgatagcctgtttgaagctgtctacaccc 917
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 gcttcaagcctctggtctccatctctcctaatgatagcctgtttgaagctgtctacaccc 180

Qy     918 tcatcaagaaccggatccatcgcctgcctgttcttgacccggtgtcaggcaacgtactcc 977
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 tcatcaagaaccggatccatcgcctgcctgttcttgacccggtgtcaggcaacgtactcc 240

Qy     978 acatcctcacacacaaacgcctgctcaagttcctgcacatctttggttccctgctgcccc 1037
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 acatcctcacacacaaacgcctgctcaagttcctgcacatctttggttccctgctgcccc 300

Qy    1038 ggccctccttcctctaccgcactatccaagatttgggcatcggcacattccgagacttgg 1097
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 ggccctccttcctctaccgcactatccaagatttgggcatcggcacattccgagacttgg 360

Qy    1098 ctgtggtgctggagacagcacccatcctgactgcactggacatctttgtggaccggcgtg 1157
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 ctgtggtgctggagacagcacccatcctgactgcactggacatctttgtggaccggcgtg 420

Qy    1158 tgtctgcactgcctgtggtcaacgaatgtggt 1189
           |||||||||| ||||||||||||||||||||
Db     421 tgtctgcactggctgtggtcaacgaatgtggt 452


RESULT   8
AAH43682
ID   AAH43682 standard; DNA; 989 BP.
XX
AC   AAH43682;
XX
DT   21-JAN-2002  (first entry)
XX
DE   PRKAG3 intron 2 - intron 4.
XX
KW   Human; AMP-activated protein kinase gamma 3 subunit; PRKAG3; variant;
KW   metabolic disease; diabetes; obesity; substitution; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   intron          1..21
FT                   /*tag= a
FT                   /number = "Intron 2"
FT                   /note= "3' portion of intron 2"
FT   exon            22..177
FT                   /*tag= b
FT                   /number= "Exon 3"
FT   intron          178..541
FT                   /*tag= c
FT                   /number= "Intron 3"
FT   exon            542..945
FT                   /*tag= d
FT                   /number= "Exon 4"
FT   intron          946..989
FT                   /*tag= e
FT                   /number= "Intron 4"
FT                   /note= "5' portion of intron 4"
XX
PN   WO200177305-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-SE00765.
XX
PR   07-APR-2000; 2000US-195665P.
XX
PA   (AREX-) AREXIS AB.
XX
PI   Andersson L,  Luthman H,  Marklund S;
XX
DR   WPI; 2001-657170/75.
XX
PT   New variants of human AMP-activated protein kinase gamma3 subunit
PT   associated with a metabolic disease e.g. diabetes or obesity and method
PT   for determining a risk estimate of diseases in subject by detecting the
PT   variant -
XX
PS   Example 1; Fig 2; 25pp; English.
XX
CC   The sequences given in AAH43681-84 represents genomic fragments
CC   encoding the human AMP-activated protein kinase gamma 3 subunit
CC   (PRKAG3). Detecting the presence of the PRKAG3 DNA, or a variant,
CC   is useful in determining a risk estimate of a metabolic disease,
CC   such as diabetes or obesity, in a subject.  The variation may occur
CC   in exons 3, 4 or 10.  In exon 3 variation may be a substitution of
CC   a G for a C at nucleotide 320, resulting in the amino acid
CC   substitution P71A; in exon 4 variation may be a substitution of a
CC   T for a C at nucleotide 550; and in exon 10 variation may be a
CC   substitution of a T for a C at nucleotide 1037, resulting in the
CC   amino acid substitution R340W.  There may also be nucleotide variation
CC   in intron 6.
XX
SQ   Sequence 989 BP; 229 A; 306 C; 286 G; 168 T; 0 other;


  Query Match             24.7%;  Score 407.2;  DB 22;  Length 989;
  Best Local Similarity   99.3%;  Pred. No. 2.5e-96;
  Matches  409;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy     247 aggtccccggtccaggccagctgctgagtccaccgggctggaggccacattccccaagac 306
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     540 aggtccccggtccaggccagctgctgagtccaccgggctggaggccacattccccaagac 599

Qy     307 cacacccttggctcaagctgatcctgccggggtgggcactccaccaacagggtgggactg 366
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     600 cacacccttggctcaagctgatcctgccggggtgggcactccaccaacagggtgggactg 659

Qy     367 cctcccctctgactgtacagcctcagctgcaggctccagcacagatgatgtggagctggc 426
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     660 cctcccctctgactgtacagcctcagctgcaggctccagcacagatgatgtggagctggc 719

Qy     427 cacggagttcccagccacagaggcctgggagtgtgagctagaaggcctgctggaagagag 486
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     720 cacggagttcccagccacagaggcctgggagtgtgagctagaaggcctgctggaagagag 779

Qy     487 gcctgccctgtgcctgtccccgcaggccccatttcccaagctgggctgggatgacgaact 546
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     780 gcctgccctgtgcctgtccccgcaggccccatttcccaagctgggctgggatgacgaact 839

Qy     547 gcggaaacccggcgcccagatctacatgcgcttcatgcaggagcacacctgctacgatgc 606
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     840 gcggaaacccggcgcccagatctacatgcgcttcatgcaggagcacacctgctacgatgc 899

Qy     607 catggcaactagctccaagctagtcatcttcgacaccatgctggagatcaag 658
           |||||||||||||||||||||||||||||||||||||||||||||||| | | |
Db     900 catggcaactagctccaagctagtcatcttcgacaccatgctggaggtgagg 951


RESULT   9
```

```
AAC98774
ID   AAC98774 standard; cDNA; 1691 BP.
XX
AC   AAC98774;
XX
DT   09-MAR-2001  (first entry)
XX
DE   Human pancreatic cancer antigen nucleotide sequence SEQ ID NO:2.
XX
KW   Human; pancreas; pancreatic cancer; pancreatic cancer antigen;
KW   detection; diagnosis; identification; cytostatic; neuroprotective;
KW   nootropic; immunomodulatory; relaxant; contraceptive; gynaecological;
KW   antiinflammatory; cardiant; gene therapy; chromosome mapping;
KW   linkage analysis; tissue identification; tissue typing; forensic;
KW   neural; immune system; muscular; reproductive; gastrointestinal;
KW   pulmonary; cardiovascular; renal; proliferative; ss.
XX
OS   Homo sapiens.
XX
PN   WO200055320-A1.
XX
PD   21-SEP-2000.
XX
PF   08-MAR-2000; 2000WO-US05989.
XX
PR   12-MAR-1999;   99US-0124270.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
XX
PI   Rosen CA,  Ruben SM;
XX
DR   WPI; 2000-579444/54.
DR   P-PSDB; AAB54009.
XX
PT   New nucleic acid that is a pancreatic cancer antigen for preventing,
PT   treating, or ameliorating a medical condition, particular pancreatic
PT   cancer, or for use in assays for diagnosing a pathological condition -
XX
PS   Claim 1; Page 498-499; 1379pp; English.
XX
CC   AAC98773 to AAC99231 encode the human pancreatic cancer associated
CC   proteins, called pancreatic cancer antigens, given in AAB54008 to
CC   AAB54466. The human pancreatic cancer antigens have cytostatic,
CC   neuroprotective, nootropic, immunomodulatory, relaxant, contraceptive,
CC   gynaecological, cardiant and antiinflammatory activities, and can be used
CC   in gene therapy. The polynucleotide and proteins can be used for
CC   preventing, treating, or ameliorating a medical condition or in assays
CC   for diagnosing a pathological condition or a susceptibility to one in a
CC   subject. Binding partners to the proteins and the activity of the
CC   proteins can be identified. The pancreatic cancer antigens can be used to
CC   detect, treat or prevent pancreatic disorders, especially cancer.
CC   Agonists and antagonists to the antigens can be screened for. The
CC   pancreatic cancer antigen polynucleotides can be used to design nucleic
CC   acid hybridisation probes that can be used in chromosome mapping, linkage
CC   analysis, tissue identification and/or typing and a variety of forensic
CC   and diagnostic methods. The proteins can be used to generate antibodies
CC   which are used to purify, detect and target the polypeptides, including
CC   both in vivo and in vitro diagnostic and therapeutic methods. The
CC   proteins can be used to treat or prevent neural, immune system, muscular,
CC   reproductive, gastrointestinal, pulmonary, cardiovascular, renal or
CC   proliferative disorders. AAC99232 to AAC99240 and AAB54467 represent
CC   sequences used in the exemplification of the present invention.
XX
SQ   Sequence 1691 BP; 458 A; 382 C; 401 G; 449 T; 1 other;

  Query Match             23.9%;  Score 393.4;  DB 21;  Length 1691;
  Best Local Similarity   65.1%;  Pred. No. 1.2e-92;
  Matches  580;  Conservative    0;  Mismatches  311;  Indels    0;  Gaps    0;

Qy     567 tctacatgcgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagc 626
           | || |    ||||||| ||   ||   |||||| ||    ||  | || ||||||||
Db     123 tgtatacttccttcatgaagtctcatcgctgctatgacctgattcccacaagctccaaat 182

Qy     627 tagtcatcttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacg 686
           | ||  | ||  || ||    ||| || | ||||| || || |||||| ||||| | ||||
Db     183 tggttgtatttgatacgtccctgcaggtgaagaaagctttttttgctttggtgactaacg 242

Qy     687 gtgtgcgggcagcccctctatgggacagcaagaagcagagctttgtggggatgctgacca 746
           |||| || || |||||| ||||||| || |||||||| || |||||||| ||||||||||
Db     243 gtgtacgagctgccccttatgggatagtaagaagcaaagttttgtgggcatgctgacca 302

Qy     747 tcactgacttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatg 806
           ||||||| ||||||    | ||||| |||||||| |  ||  || |||| ||||||||||
Db     303 tcactgatttcatcaatatcctgcaccgctactataaatcagccttggtacagatctatg 362

Qy     807 agattgaacaacataagattgagacctggagggagatctacctgcaaggctgcttcaagc 866
           || | ||| |||| ||||| || || ||||| ||| | || || || | || ||| || |
Db     363 agctagaagaacacaagatagaaacttggagagaggtgtatctccaggactcctttaaac 422

Qy     867 ctctggtctccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaaga 926
           | || |||| ||| |||||||||| ||| ||||||| |||||||   |  | ||   ||
Db     423 cgcttgtctgcatttctcctaatgccagcttgtttgatgctgtctcttcattaattcgga 482

Qy     927 accggatccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctca 986
           ||  |||||| | ||||| ||| ||||||| |  |||||||     |  |||||||||
Db     483 acaagatccacaggctgccagttattgacccagaatcaggcaatactttgtacatcctca 542

Qy     987 cacacaaacgcctgctcaagttcctgcacatctttggttccctgctgccccggccctcct 1046
           | ||||| ||| | || ||||||||  |  | |||    |   | | |||  |||    |
Db     543 cccacaagcgcattctgaagttcctcaaattgtttatcactgagttccccaagccagagt 602

Qy    1047 tcctctaccgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgc 1106
           || | | |    || |  |||| |    || ||||| |       |  | ||| |||| |
Db     603 tcatgtccaagtctctggaagagctacagattggcacctatgccaatattgctatggttc 662

Qy    1107 tggagacagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcac 1166
              ||  | ||| ||      || ||||  || |||||  | |  || || || || |
Db     663 gcactaccacccccgtctatgtggctctggggatttttgtacagcatcgagtctcagccc 722

Qy    1167 tgcctgtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattc 1226
           |||| |||||  | ||    || |  || |||| | |||| |||   |||||||| ||
Db     723 tgccagtggtggatgagaaggggcgtgtggtggacatctactccaagtttgatgttatca 782

Qy    1227 acctggctgcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggc 1286
           | ||||| ||  |  | |||||||| |||| ||  |   ||||   ||||| ||   |
Db     783 atctggcagcagaaaagacctacaacaacctagatgtatctgtgactaaagccttgcaac 842

Qy    1287 agaggacactatgtctggagggagtcctttcctgccagccccacgagagcttgggggaag 1346
           |  |  |||  |   | ||||| || ||    ||| | |  || |||   ||| |
Db     843 atcgatcacattactttgagggtgttctcaagtgctacctgcatgagactctggagacca 902

Qy    1347 tgatcgacaggattgctcgggagcaggtacacaggctggtgctagtggacgagacccagc 1406
           | ||| ||||| | |     |  |||| ||| | || ||  | ||||| || |   |
Db     903 tcatcaacaggctagtggaagcagaggttcaccgacttgtagtggtggatgaaaatgatg 962

Qy    1407 atctcttgggcgtggtctccctctccgacatccttcaggcactggtgctca 1457
              ||  |||  | || || || || |||||||| ||||| ||||||||||
Db     963 tggtcaagggaattgtatcactgtctgacatcctgcaggccctggtgctca 1013

RESULT  10
AAT85927
ID   AAT85927 standard; DNA; 1576 BP.
XX
AC   AAT85927;
XX
DT   16-MAR-1998  (first entry)
XX
DE   Mammalian AMPK-gamma subunit DNA.
XX
KW   5'-AMP activated protein kinase; AMPK; non-catalytic subunit;
KW   protein phosphorylation; cholesterol; fatty acid; gamma subunit;
KW   hormone sensitive lipase; HSL; ss.
XX
OS   Mammalia.
```

```
XX
PN   WO9725341-A1.
XX
PD   17-JUL-1997.
XX
PF   07-JAN-1997;   97WO-US00270.
XX
PR   08-JAN-1996;   96AU-0007450.
XX
PA   (DART-) DARTMOUTH COLLEGE.
PA   (SVIN-) ST VINCENTS INST MEDICAL RES.
XX
PI   Kemp BE,  Mitchelhill KI,  Stapleton DI,  Witters LA;
XX
DR   WPI; 1997-372811/34.
XX
PT   New isolated 5'-AMP-activated protein kinase subunit(s) - used to
PT   develop products for treating e.g. hyper-cholesterolaemia, obesity,
PT   hypoxia, ischaemia, nutrition disorders or diabetes mellitus
XX
PS   Claim 16; Page 53; 63pp; English.
XX
CC   This polynucleotide sequence encodes a mammalian 5'-AMP-activated
CC   protein kinase (AMPK) non-catalytic gamma subunit. AMPK polypeptides can
CC   be used to identify compounds which regulate the action of kinases. Such
CC   fragments can be used to reduce biosynthesis of cholesterol and fatty
CC   acids. They may also be used to inhibit the release of these molecules
CC   from intracellular stores by hormone sensitive lipase (HSL). They may
CC   also be used to reduce cellular malonyl CoA levels and promote the
CC   beta-oxidation of fatty acids by mitochondria. AMPK-alphal fragments
CC   could be used in the treatment of e.g. hypercholesterolaemia,
CC   hyperlipidaemia, obesity, clinical syndromes associated with hypoxia or
CC   ischaemia (e.g. myocardial infarction) disorders of nutrition and
CC   diabetes mellitus.
XX
SQ   Sequence 1576 BP; 401 A; 375 C; 378 G; 422 T; 0 other;

  Query Match             23.7%;  Score 390.2;  DB 18;  Length 1576;
  Best Local Similarity   64.9%;  Pred. No. 8.1e-92;
  Matches  578;  Conservative    0;  Mismatches  313;  Indels    0;  Gaps    0;

Qy     567 tctacatgcgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagc 626
           | || |    ||||||| ||   ||   |||||| ||    ||  | || ||||||||
Db     176 tgtatacttccttcatgaagtctcatcgctgctatgacctgattcccacaagctccaaat 235

Qy     627 tagtcatcttcgacaccatgctggagatcaagaaggcctttctttgctctggtggccaacg 686
           | ||  | || || ||    ||| || | ||||| || || |||||| ||||| | ||||
Db     236 tggttgtatttgatacgtccctgcaggtgaagaaagctttttttgctttggtgactaacg 295

Qy     687 gtgtgcgggcagcccctctatgggacagcaagaagcagagctttgtggggatgctgacca 746
           |||| || || |||||| ||||||| || |||||||| || |||||||| ||||||||||
Db     296 gtgtacgagctgccccttatgggatagtaagaagcaaagttttgtgggcatgctgacca 355

Qy     747 tcactgacttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatg 806
           ||||||| ||||||    | ||||| |||||||| |  ||  |  |||| ||||||||||
Db     356 tcactgatttcatcaatatcctgcaccgctactataaatcagcgttggtacagatctatg 415

Qy     807 agattgaacaacataagattgagacctggagggagatctacctgcaaggctgcttcaagc 866
           || | ||| |||| ||||| || || ||||| ||| | || || || | || ||| || |
Db     416 agctagaagaacacaagatagaaacttggagagaggtgtatctccaggactcctttaaac 475

Qy     867 ctctggtctccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaaga 926
           | || |||| ||| |||||||||||  ||| ||||||| |||||||   |  | ||    |
Db     476 cgcttgtctgcatttctcctaatgccagcttgtttgatgctgtctcttcattaattcgca 535

Qy     927 accggatccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctca 986
           ||  ||||||  | ||||| ||| |||||||  |  ||||||||     |  |||||||||
Db     536 acaagatccacaggctgccagttattgacccagaatcaggcaatactttgtacatcctca 595

Qy     987 cacacaaacgcctgctcaagttcctgcacatctttggttccctgctgccccggccctcct 1046
           | ||||| ||| | || ||||||||  |  | |||    |   | | |||  |||    |
Db     596 cccacaagcgcattctgaagttcctcaaattgtttatcactgagttccccaagccagagt 655

Qy    1047 tcctctaccgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgc 1106
           || | | |    || |  |||| |    || ||||| |      |  | ||| |||| |
Db     656 tcatgtccaagtctctggaagagctacagattggcacctatgccaatattgctatggttc 715

Qy    1107 tggagacagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcac 1166
              ||  | ||| ||      || ||||  || |||||  | |  || || || || |
Db     716 gcactaccacccccgtctatgtggctctggggatttttgtacagcatcgagtctcagccc 775

Qy    1167 tgcctgtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattc 1226
           |||| |||||  | ||    || |  || |||| | |||| |||   |||||||| ||
Db     776 tgccagtggtggatgagaaggggcgtgtggtggacatctactccaagtttgatgttatca 835

Qy    1227 acctggctgcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggc 1286
           | ||||| ||  |  | |||||||||| |||| ||  |   ||||    ||||| ||   |
Db     836 atctggcagcagaaaagacctacaacaacctagatgtatatgtgactaaagccttgcaac 895

Qy    1287 agaggacactatgtctggagggagtccttttcctgccagccccacgagagcttgggggaag 1346
           |  |  |||  |   | ||||| || ||    ||| | |  || ||||   ||| |
Db     896 atcgatcacattactttgagggtgttctcaagtgctacctgcatgagactctggagacca 955

Qy    1347 tgatcgacaggattgctcgggagcaggtacacaggctggtgctagtggacgagacccagc 1406
           | ||| ||||| | |     |   |||| ||| | || ||  | ||||| || |   |
Db     956 tcatcaacaggctagtggaagcagaggttcaccgacttgtagtggtggatgaaaatgatg 1015

Qy    1407 atctcttgggcgtggtctccctctccgacatccttcaggcactggtgctca 1457
              ||  |||  | || || || || |||||||| ||||| ||||||||||
Db    1016 tggtcaagggaattgtatcactgtctgacatcctgcaggccctggtgctca 1066


RESULT  11
AAX06882
ID   AAX06882 standard; cDNA; 1435 BP.
XX
AC   AAX06882;
XX
DT   26-APR-1999  (first entry)
XX
DE   Disease associated protein kinase DAPK-7 cDNA.
XX
KW   DAPK-7; disease associated protein kinase; human; diagnosis;
KW   therapy; adult respiratory distress syndrome; allergy; asthma;
KW   arteriosclerosis; bronchitis; emphysema; hypereosinophilia;
KW   myocardial inflammation; pericardial inflammation; anaemia;,
KW   rheumatoid arthritis; Addison's disease; AIDS; atherosclerosis;
KW   atopic dermatitis; dermatomyositis; diabetes mellitus;
KW   glomerulonephritis; gout; Grave's disease; lupus erythematosus;
KW   multiple sclerosis; myasthenia gravis; osteoarthritis;
KW   osteoporosis; pancreatitis; polycystic kidney disease;
KW   polymyositis; scleroderma; Sjorgren's syndrome;
KW   autoimmune thyroiditis; cancer; infection; trauma;
KW   cell proliferation; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             265..1503
FT                   /*tag=  a
XX
PN   WO9858052-A2.
XX
PD   23-DEC-1998.
XX
PF   19-JUN-1998;   98WO-US12813.
XX
PR   19-JUN-1997;   97US-0878989.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Bandman O,  Corley NC,  Goli SK,  Guegler KJ,  Hillman JL;
PI   Lal P,  Shah P;
XX
DR   WPI; 1999-080952/07.
```

```
DR   P-PSDB; AAW88438.
XX
PT   New disease associated protein kinases - used to stimulate cell
PT   proliferation and to treat the immune response and cancer
XX
PS   Claim 5; Page 66-67; 93pp; English.
XX
CC   This cDNA sequence codes for human disease associated protein
CC   kinase DAPK-7 (see AAW88438).  DAPK-7 cDNA was first identified in
CC   the PENITUT01 cDNA library using a computer search for amino acid
CC   alignments, and a consensus sequence was derived from the extended
CC   and overlapping Incyte clones 3075712/HEARNOT01, 842220/PROSTUT05,
CC   1364747/SCORNON02, 145972 and 145802/PENITUT01 and 1479332/CORPNOT02.
CC   DAPK-7 shows 73% homology with the human foetal liver AMPK gamma
CC   subunit (GI 1335856), and is associated with cDNA libraries which
CC   are immortalised or cancerous and show inflammatory or immune
CC   responses. The invention provides disease associated protein kinases
CC   DAPK-1 to DAPK-7 (see AAW88432-38) and cDNA clones encoding them (see
CC   AAX06831-36 and AAX06882), as well as expression vectors, host cells,
CC   agonists, antagonists and antibodies. The invention further provides
CC   uses of such products in the diagnosis, prevention and treatment of
CC   diseases associated with cell proliferation, especially cancer or an
CC   immune response (claimed).  Conditions that may be treated include
CC   adult respiratory distress syndrome, allergies, asthma,
CC   arteriosclerosis, bronchitis, emphysema, hypereosinophilia,
CC   myocardial or pericardial inflammation, rheumatoid arthritis,
CC   Addison's disease, AIDS, anaemia, atherosclerosis, various diseases
CC   of the digestive system, atopic dermatitis, dermatomyositis, diabetes
CC   mellitus, glomerulonephritis, gout, Grave's disease, lupus
CC   erythematosus, multiple sclerosis, myasthenia gravis,
CC   osteoarthritis, osteoporosis, pancreatitis, polycystic kidney
CC   disease, polymyositis, scleroderma, Sjorgren's syndrome, autoimmune
CC   thyroiditis, complications of cancer, extracorporeal circulation,
CC   viral, bacterial, fungal, parasitic, protozoal and helminthic
CC   infections, and trauma (disclosed).  The DAPK nucleic acids are
CC   also used in a method for detection of DAPK expression levels in a
CC   biological sample.
XX
SQ   Sequence 1435 BP; 421 A; 298 C; 331 G; 385 T; 0 other;

  Query Match             22.3%;  Score 366.6;  DB 20;  Length 1435;
  Best Local Similarity   62.9%;  Pred. No. 1.1e-85;
  Matches  567;  Conservative    0;  Mismatches  334;  Indels    0;  Gaps    0;

Qy     567 tctacatgcgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagc 626
           | ||||||||| ||||||  |   ||||  || || ||   | |  |||| || || ||||
Db     183 tttacatgcgattcatgaggtcacacaagtgttatgacatcgttccaaccagttcaaagc 242

Qy     627 tagtcatcttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacg 686
           | ||  |||| || || |   |  |  | || ||||||||||||||| |||| |||||||
Db     243 ttgttgtctttgatactacattacaagttaaaaaggccttctttgctttggtagccaacg 302

Qy     687 gtgtgcgggcagcccctctatgggacagcaagaagcagagctttgtggggatgctgacca 746
           |||| || ||||| || || ||||| || || || || || ||||| || ||||| || |
Db     303 gtgtccgagcagcgccactgtgggagagtaaaaaacaaagttttgtaggaatgctaacaa 362

Qy     747 tcactgacttcatcctggtgctgcatcgctactacaggtccccctggtccagatctatg 806
           | || || |||||     | || ||| | ||||| |  || ||  |||| ||||| ||||
Db     363 ttacagatttcataaatatactacatagatactataaatcacctatggtacagatttatg 422

Qy     807 agattgaacaacataagattgagacctggagggagatctacctgcaaggctgcttcaagc 866
           |  | ||  |||||| ||||| || |||||||| | ||  | ||||     || ||||
Db     423 aattagaggaacataaaattgaaacatggagggagctttatttacaagaaacatttaagc 482

Qy     867 ctctggtctccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaaga 926
           || | ||    || |||||  |||  ||||| || || ||||| ||| || | ||||| |
Db     483 ctttagtgaatatatctccagatgcaagcctcttcgatgctgtatactccttgatcaaaa 542

Qy     927 accggatccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctca 986
           |    |||||  |  |||| ||| |||||||  |    || || | |||  | || || |
Db     543 ataaaatccacagattgcccgttattgaccctatcagtgggaatgcactttatatactta 602

Qy     987 cacacaaacgcctgctcaagttcctgcacatctttggttccctgctgccccggccctcct 1046
           | |||||| |  | ||||||||||| ||  | |||   ||     ||||  |||  |||
Db     603 cccacaaaagaatcctcaagttcctccagcttttatgtctgatatgccaaagcctgcct 662

Qy    1047 tcctctaccgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgc 1106
           || |  | |  |   |  | ||  | || || || || | ||   || | ||  |  | |
Db     663 tcatgaagcagaacctggatgagcttggaataggaacgtaccacaacattgccttcatac 722

Qy    1107 tggagacagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcac 1166
                | |||||| | |  ||  || |||| ||||||||  | ||  | || || | |
Db     723 atccagacactcccatcatcaaagccttgaacatatttgtggaaagacgaatatcagctc 782

Qy    1167 tgcctgtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattc 1226
           |||||| ||  | || |  ||  | || || |   | ||||||  |||||||| |||
Db     783 tgcctgttgtggatgagtcaggaaaagttgtagatatttattccaaatttgatgtaatta 842

Qy    1227 acctggctgcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggc 1286
           | || |||||  || |||| |||||  |||| || || |  |||    | |||||  ||
Db     843 atcttgctgctgagaaaacatacaataacctagatatcacggtgacccaggcccttcagc 902

Qy    1287 agaggacactatgtctggagggagtcctttcctgccagccccacgagagcttgggggaag 1346
           |  |  |||  | | | || || ||  |    ||| |    |  || |  ||| |
Db     903 accgttcacagtattttgaaggtgttgtgaagtgcaataagctggaaatactggagacca 962

Qy    1347 tgatcgacaggattgctcgggagcaggtacacaggctggtgctagtggacgagacccagc 1406
           |  | ||||| || |   | |   |||| ||  |||||||| | ||  | ||  |  |
Db     963 tcgtggacagaatagtaagagctgaggtccatcggctggtggtggtaaatgaagcagata 1022

Qy    1407 atctcttgggcgtggtctccctctccgacatccttcaggcactggtgctcagccctgctg 1466
            | |  |||| |  | ||||| || ||||| || || || ||| | ||||  || || |
Db    1023 gtattgtgggtattatttccctgtcggacattctgcaagccctgatcctcacaccagcag 1082

Qy    1467 g 1467
           |
Db    1083 g 1083


RESULT  12
AAH14839
ID   AAH14839 standard; cDNA; 2223 BP.
XX
AC   AAH14839;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:12660.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
```

PT full-length cDNAs -
XX
PS Claim 8; SEQ ID 12660; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesising 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 2223 BP; 633 A; 431 C; 496 G; 663 T; 0 other;

Query Match 22.3%; Score 366.6; DB 22; Length 2223;
Best Local Similarity 62.9%; Pred. No. 1.3e-85;
Matches 567; Conservative 0; Mismatches 334; Indels 0; Gaps 0;

```
Qy    567 tctacatgcgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagc 626
          | ||||||| |||||| |   |||| || || ||   | |  |||| || || ||||
Db    215 tttacatgcgattcatgaggtcacacaagtgttatgacatcgttccaaccagttcaaagc 274

Qy    627 tagtcatcttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacg 686
          | ||  |||| || || |   |  |  | || ||||||||||||||| |||| |||||||
Db    275 ttgttgtctttgatactacattacaagttaaaaaggccttctttgctttggtagccaacg 334

Qy    687 gtgtgcgggcagcccctctatgggacagcaagaagcagagctttgtggggatgctgacca 746
          |||| || ||||| || || ||||| || || || || || ||||| || ||||| || |
Db    335 gtgtccgagcagcgccactgtgggagagtaaaaaacaaagttttgtaggaatgctaacaa 394

Qy    747 tcactgacttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatg 806
          | || || |||||     | || ||| | ||||| |  || ||  |||| ||||| ||||
Db    395 ttacagatttcataaatatactacatagatactataaatcacctatggtacagatttatg 454

Qy    807 agattgaacaacataagattgagacctggagggagatctacctgcaaggctgcttcaagc 866
          |  | ||  ||||||| ||||| || |||||||||| | ||  | ||||     || ||||
Db    455 aattagaggaacataaaattgaaacatggagggagctttatttacaagaaacatttaagc 514

Qy    867 ctctggtctccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaaga 926
          || | ||    || |||||  |||  ||||| || || ||||| ||| || | ||||| |
Db    515 ctttagtgaatatatctccagatgcaagcctcttcgatgctgtatactccttgatcaaaa 574

Qy    927 accggatccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctca 986
          |    |||||  |  |||| ||| |||||||  |    || || | |||  | || || |
Db    575 ataaaatccacagattgcccgttattgaccctatcagtgggaatgcactttatatactta 634

Qy    987 cacacaaacgcctgctcaagttcctgcacatctttggttccctgctgccccggccctcct 1046
          | |||||| |  | ||||||||||| ||  | |||   ||     ||||   |||  |||
Db    635 cccacaaaagaatcctcaagttcctccagcttttatgtctgatatgccaaagcctgcct 694

Qy   1047 tcctctaccgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgc 1106
          || |  | | |   |  | ||  | || || || || | ||   || | ||  |  | |
Db    695 tcatgaagcagaacctggatgagcttggaataggaacgtaccacaacattgccttcatac 754

Qy   1107 tggagacagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcac 1166
                | |||||| | |  ||  || |||| ||||||||  | ||  | || || | |
Db    755 atccagacactcccatcatcaaagccttgaacatatttgtggaaagacgaatatcagctc 814

Qy   1167 tgcctgtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattc 1226
          ||||||| ||  | || |  ||  | || || |   | ||||||   |||||||| |||
Db    815 tgcctgttgtggatgagtcaggaaaagttgtagatatttattccaaatttgatgtaatta 874

Qy   1227 acctggctgcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggc 1286
          | || |||||  || |||| |||||  |||| || || |  |||    | |||||   ||
Db    875 atcttgctgctgagaaaacatacaataacctagatatcacggtgacccaggcccttcagc 934

Qy   1287 agaggacactatgtctggagggagtcctttcctgccagccccacgagagcttgggggaag 1346
          |  |  |||  | | | || || ||  |    ||| |    |  || |   ||| |
Db    935 accgttcacagtattttgaaggtgttgtgaagtgcaataagctggaaatactggagacca 994

Qy   1347 tgatcgacaggattgctcgggagcaggtacacaggctggtgctagtggacgagacccagc 1406
          |  | ||||| || |   | |   |||| ||  |||||||| | ||  | ||  |  |
Db    995 tcgtggacagaatagtaagagctgaggtccatcggctggtggtggtaaatgaagcagata 1054

Qy   1407 atctcttgggcgtggtctccctctccgacatccttcaggcactggtgctcagccctgctg 1466
           | |  |||| |  | ||||| || ||||| || || || ||| | ||||  || || |
Db   1055 gtattgtgggtattatttccctgtcggacattctgcaagccctgatcctcacgccagcag 1114

Qy   1467 g 1467
          |
Db   1115 g 1115
```

RESULT 13
ABL18857
ID ABL18857 standard; DNA; 3261 BP.
XX
AC ABL18857;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 8044.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 8044; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

```
XX
SQ   Sequence 3261 BP; 810 A; 952 C; 931 G; 568 T; 0 other;

  Query Match            18.2%;  Score 299.2;  DB 23;  Length 3261;
  Best Local Similarity  58.6%;  Pred. No. 6.1e-68;
  Matches  539;  Conservative    0;  Mismatches  378;  Indels    3;  Gaps    1;

Qy     542 gaactgcggaaacccggcgcccagatctacatgcgcttcatgcaggagcacacctgctac 601
           || |||     |   || | | ||||||| | ||   ||| | |     |||| ||||| 
Db    1736 gatctggaagaggacgactcacagatcttcgtgaagttctttcgttttcacaagtgctat 1795

Qy     602 gatgccatggcaactagctccaagctagtcatcttcgacaccatgctggagatcaagaag 661
           |||   ||  | ||   | ||||| | ||  ||||||||||||  |||     | ||||||
Db    1796 gatctgatacccacctccgccaagttggttgtcttcgacacccagcttcttgtaaagaag 1855

Qy     662 gccttctttgctctggtggccaacggtgtgcgggcagcccctctatgggacagcaagaag 721
           |||||||  || || ||  |||||||||||| || || || || ||||||     ||||| 
Db    1856 gccttctacgccctcgtctacaacggtgtgcgagcggcaccgctctgggattcggagaag 1915

Qy     722 cagagctttgtggggatgctgaccatcactgacttcatcctggtgctgcatcgctactac 781
           ||    || ||||| ||||| |||||||| ||||| |||  | | |||||     || |||
Db    1916 caacagttcgtgggcatgctaaccatcacggactttatcaagatcctgcaaatgtattac 1975

Qy     782 aggtccccccctggtccagatctatgagattgaacaacataagattgagacctggagggag 841
           |  || ||    |     ||  |  || | ||| | || ||  | || || ||| |   |
Db    1976 aaatcgccaaatgcgtccatggagcagctggaagagcacaaactggacacgtggcg---g 2032

Qy     842 atctacctgcaaggctgcttcaagcctctggtctccatctctcctaatgatagcctgttt 901
           | |   |||||   |    | | |||  |||||  ||||   ||  |||    ||| |
Db    2033 agcgtgctgcacaaccaggtgatgccgttggtcagcatcggaccggatgcgtccctctac 2092

Qy     902 gaagctgtctacaccctcatcaagaaccggatccatcgcctgcctgttcttgacccggtg 961
           || ||  || | |  |||||| | | ||| || |||||||||||  ||  | || |||| |
Db    2093 gatgccatcaaaattctcatccacagccgcatacatcgcctgcccgtcatcgatccggcg 2152

Qy     962 tcaggcaacgtactccacatcctcacacacaaacgcctgctcaagttcctgcacatcttt 1021
            | |||||| || ||| ||||||| ||||| |||||| | || | ||||||   | | |
Db    2153 accggcaatgtcctctacatcctgacacataaacgcatacttaggttccttttcctatac 2212

Qy    1022 ggttccctgctgccccggccctccttcctctaccgcactatccaagatttgggcatcggc 1081
             |       | ||   |||| | | | |   |   | | | |  ||  ||   || |||
Db    2213 attaatgaattaccaaagcccgcgtacatgcaaaaaagtttgcgcgaactgaagattggc 2272

Qy    1082 acattccgagacttggctgtggtgctggagacagcacccatcctgactgcactggacatc 1141
           || |      || | |     |     |||||  |   |||| | || || ||  | |
Db    2273 acctataacaacatcgagaccgccgacgagacgacgagcatcatcacggcgctcaagaaa 2332

Qy    1142 tttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaatgtggtcaggtcgtgggc 1201
           |||||||| || || || || || ||||||  | ||  |      |||||   |||| | |
Db    2333 tttgtggagcgacgagtctcagccctgccactagtggattccgatggtcgcctcgttgac 2392

Qy    1202 ctctattcccgctttgatgtgattcacctggctgcccagcaaacctacaaccacctggac 1261
            | ||  |    ||||||||||||| | || || ||| || |||||||||||| | || ||
Db    2393 atttacgcaaagtttgatgtgattaatctcgccgccgagaaaacctacaacgatctcgat 2452

Qy    1262 atgagtgtgggagaagccctgaggcagaggacactatgtctggagggagtcctttcctgc 1321
            |     || |  |||||     |||  |||     ||  | ||||| || |     |||
Db    2453 gtttcgctgcgcaaagccaacgagcaccggaacgagtggttcgagggcgtgcagaagtgc 2512

Qy    1322 cagccccacgagagcttgggggaagtgatcgacaggattgctcgggagcaggtacacagg 1381
            | |   ||||     |         | || ||  | || |  || |   | |||||  |
Db    2513 aatctggacgaatcgctctacacgatcatggaacgaatcgtccgcgccgaagtacatcga 2572

Qy    1382 ctggtgctagtggacgagacccagcatctcttgggcgtggtctccctctccgacatcctt 1441
           ||||||  | || |||||||   |   |  | ||| |  |||| || ||||| || ||
Db    2573 ctggtggtggtcgacgagaatcgcaaagtgatcggcattatctcgctgtccgatatactg 2632

Qy    1442 caggcactggtgctcagccc 1461
           |     || || ||  | ||
Db    2633 ctctacctcgtcctgcgacc 2652
```

```
RESULT  14
AAH43684
ID   AAH43684 standard; DNA; 1014 BP.
XX
AC   AAH43684;
XX
DT   21-JAN-2002  (first entry)
XX
DE   PRKAG3 intron 10 - 3'UTR.
XX
KW   Human; AMP-activated protein kinase gamma 3 subunit; PRKAG3; variant;
KW   metabolic disease; diabetes; obesity; substitution; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   intron          1..41
FT                   /*tag= a
FT                   /number = "Intron 10"
FT                   /note= "3' portion of intron 10"
FT   exon            42..79
FT                   /*tag= b
FT                   /number= "Exon 11"
FT   intron          80..249
FT                   /*tag= c
FT                   /number= "Intron 11"
FT   exon            250..396
FT                   /*tag= d
FT                   /number= "Exon 12"
FT   intron          397..739
FT                   /*tag= e
FT                   /number= "Intron 12"
FT   exon            740..856
FT                   /*tag= f
FT                   /number= "Exon 13"
FT   3'UTR           857..1014
FT                   /*tag= g
XX
PN   WO200177305-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-SE00765.
XX
PR   07-APR-2000; 2000US-195665P.
XX
PA   (AREX-) AREXIS AB.
XX
PI   Andersson L,  Luthman H,  Marklund S;
XX
DR   WPI; 2001-657170/75.
XX
PT   New variants of human AMP-activated protein kinase gamma3 subunit
PT   associated with a metabolic disease e.g. diabetes or obesity and method
PT   for determining a risk estimate of diseases in subject by detecting the
PT   variant -
XX
PS   Example 1; Fig 4; 25pp; English.
XX
CC   The sequences given in AAH43681-84 represents genomic fragments
CC   encoding the human AMP-activated protein kinase gamma 3 subunit
CC   (PRKAG3). Detecting the presence of the PRKAG3 DNA, or a variant,
CC   is useful in determining a risk estimate of a metabolic disease,
CC   such as diabetes or obesity, in a subject.  The variation may occur
CC   in exons 3, 4 or 10.  In exon 3 variation may be a substitution of
CC   a G for a C at nucleotide 320, resulting in the amino acid
CC   substitution P71A; in exon 4 variation may be a substitution of a
CC   T for a C at nucleotide 550; and in exon 10 variation may be a
CC   substitution of a T for a C at nucleotide 1037, resulting in the
CC   amino acid substitution R340W.  There may also be nucleotide variation
CC   in intron 6.
XX
SQ   Sequence 1014 BP; 192 A; 325 C; 271 G; 226 T; 0 other;
```

```
  Query Match             16.9%;  Score 278.8;  DB 22;  Length 1014;
  Best Local Similarity   97.6%;  Pred. No. 8.3e-63;
  Matches  283;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy    1358 attgctcgggagcaggtacacaggctggtgctagtggacgagacccagcatctcttgggc 1417
           |||||       ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     725 attgcggacccacaggtacacaggctggtgctagtggacgagacccagcatctcttgggc 784

Qy    1418 gtggtctccctctccgacatccttcaggcactggtgctcagccctgctggcatcgatgcc 1477
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     785 gtggtctccctctccgacatccttcaggcactggtgctcagccctgctggcatcgatgcc 844

Qy    1478 ctcggggcctgagaagatctgagtcctcaatcccaagccacctgcacacctggaagccaa 1537
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     845 ctcggggcctgagaagatctgagtcctcaatcccaagccacctgcacacctggaagccaa 904

Qy    1538 tgaagggaactggagaactcagccttcatcttcccccaccccccatttgctggttcagcta 1597
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     905 tgaagggaactggagaactcagccttcatcttcccccaccccccatttgctggttcagcta 964

Qy    1598 tgattcaggtaggctctgccctgggccatgacaccagcctcttagtcttc 1647
           |||||||||||||||||||||||||||||||||||||||||||||||||
Db     965 tgattcaggtaggctctgccctgggccatgacaccagcctcttagtcttc 1014


RESULT  15
AAK72740/c
ID   AAK72740 standard; DNA; 3425 BP.
XX
AC   AAK72740;
XX
DT   06-NOV-2001  (first entry)
XX
DE   Human immune/haematopoietic antigen genomic sequence SEQ ID NO:27552.
XX
KW   Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW   cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS   Homo sapiens.
XX
PN   WO200157182-A2.
XX
PD   09-AUG-2001.
XX
PF   17-JAN-2001; 2001WO-US01354.
XX
PR   31-JAN-2000; 2000US-0179065.
PR   04-FEB-2000; 2000US-0180628.
PR   24-FEB-2000; 2000US-0184664.
PR   02-MAR-2000; 2000US-0186350.
PR   16-MAR-2000; 2000US-0189874.
PR   17-MAR-2000; 2000US-0190076.
PR   18-APR-2000; 2000US-0198123.
PR   19-MAY-2000; 2000US-0205515.
PR   07-JUN-2000; 2000US-0209467.
PR   28-JUN-2000; 2000US-0214886.
PR   30-JUN-2000; 2000US-0215135.
PR   07-JUL-2000; 2000US-0216647.
PR   07-JUL-2000; 2000US-0216880.
PR   11-JUL-2000; 2000US-0217487.
PR   11-JUL-2000; 2000US-0217496.
PR   14-JUL-2000; 2000US-0218290.
PR   26-JUL-2000; 2000US-0220963.
PR   26-JUL-2000; 2000US-0220964.
PR   14-AUG-2000; 2000US-0224518.
PR   14-AUG-2000; 2000US-0224519.
PR   14-AUG-2000; 2000US-0225213.
PR   14-AUG-2000; 2000US-0225214.
PR   14-AUG-2000; 2000US-0225266.
PR   14-AUG-2000; 2000US-0225267.
PR   14-AUG-2000; 2000US-0225268.
PR   14-AUG-2000; 2000US-0225270.
PR   14-AUG-2000; 2000US-0225447.
PR   14-AUG-2000; 2000US-0225757.
PR   14-AUG-2000; 2000US-0225758.
PR   14-AUG-2000; 2000US-0225759.
PR   18-AUG-2000; 2000US-0226279.
PR   22-AUG-2000; 2000US-0226681.
PR   22-AUG-2000; 2000US-0226868.
PR   22-AUG-2000; 2000US-0227182.
PR   23-AUG-2000; 2000US-0227009.
PR   30-AUG-2000; 2000US-0228924.
PR   01-SEP-2000; 2000US-0229287.
PR   01-SEP-2000; 2000US-0229343.
PR   01-SEP-2000; 2000US-0229344.
PR   01-SEP-2000; 2000US-0229345.
PR   05-SEP-2000; 2000US-0229509.
PR   05-SEP-2000; 2000US-0229513.
PR   06-SEP-2000; 2000US-0230437.
PR   06-SEP-2000; 2000US-0230438.
PR   08-SEP-2000; 2000US-0231242.
PR   08-SEP-2000; 2000US-0231243.
PR   08-SEP-2000; 2000US-0231244.
PR   08-SEP-2000; 2000US-0231413.
PR   08-SEP-2000; 2000US-0231414.
PR   08-SEP-2000; 2000US-0232080.
PR   08-SEP-2000; 2000US-0232081.
PR   12-SEP-2000; 2000US-0231968.
PR   14-SEP-2000; 2000US-0232397.
PR   14-SEP-2000; 2000US-0232398.
PR   14-SEP-2000; 2000US-0232399.
PR   14-SEP-2000; 2000US-0232400.
PR   14-SEP-2000; 2000US-0232401.
PR   14-SEP-2000; 2000US-0233063.
PR   14-SEP-2000; 2000US-0233064.
PR   14-SEP-2000; 2000US-0233065.
PR   21-SEP-2000; 2000US-0234223.
PR   21-SEP-2000; 2000US-0234274.
PR   25-SEP-2000; 2000US-0234997.
PR   25-SEP-2000; 2000US-0234998.
PR   26-SEP-2000; 2000US-0235484.
PR   27-SEP-2000; 2000US-0235834.
PR   27-SEP-2000; 2000US-0235836.
PR   29-SEP-2000; 2000US-0236327.
PR   29-SEP-2000; 2000US-0236367.
PR   29-SEP-2000; 2000US-0236368.
PR   29-SEP-2000; 2000US-0236369.
PR   29-SEP-2000; 2000US-0236370.
PR   02-OCT-2000; 2000US-0236802.
PR   02-OCT-2000; 2000US-0237037.
PR   02-OCT-2000; 2000US-0237038.
PR   02-OCT-2000; 2000US-0237039.
PR   02-OCT-2000; 2000US-0237040.
PR   13-OCT-2000; 2000US-0239935.
PR   13-OCT-2000; 2000US-0239937.
PR   20-OCT-2000; 2000US-0240960.
PR   20-OCT-2000; 2000US-0241221.
PR   20-OCT-2000; 2000US-0241785.
PR   20-OCT-2000; 2000US-0241786.
PR   20-OCT-2000; 2000US-0241787.
PR   20-OCT-2000; 2000US-0241808.
PR   20-OCT-2000; 2000US-0241809.
PR   20-OCT-2000; 2000US-0241826.
PR   01-NOV-2000; 2000US-0244617.
PR   08-NOV-2000; 2000US-0246474.
PR   08-NOV-2000; 2000US-0246475.
PR   08-NOV-2000; 2000US-0246476.
PR   08-NOV-2000; 2000US-0246477.
PR   08-NOV-2000; 2000US-0246478.
PR   08-NOV-2000; 2000US-0246523.
PR   08-NOV-2000; 2000US-0246524.
PR   08-NOV-2000; 2000US-0246525.
PR   08-NOV-2000; 2000US-0246526.
PR   08-NOV-2000; 2000US-0246527.
```

PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
DR WPI; 2001-483426/52.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Disclosure; SEQ ID NO 27552; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patients own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/haematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK87694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 3425 BP; 916 A; 794 C; 951 G; 764 T; 0 other;

Query Match 15.3%; Score 252; DB 22; Length 3425;

Best Local Similarity 100.0%; Pred. No. 1.3e-55;
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy   1396 cgagacccagcatctcttgggcgtggtctccctctccgacatccttcaggcactggtgct 1455
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3425 CGAGACCCAGCATCTCTTGGGCGTGGTCTCCCTCTCCGACATCCTTCAGGCACTGGTGCT 3366

Qy   1456 cagccctgctggcatcgatgccctcggggcctgagaagatctgagtcctcaatcccaagc 1515
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3365 CAGCCCTGCTGGCATCGATGCCCTCGGGGCCTGAGAAGATCTGAGTCCTCAATCCCAAGC 3306

Qy   1516 cacctgcacacctggaagccaatgaagggaactggagaactcagccttcatcttccccca 1575
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3305 CACCTGCACACCTGGAAGCCAATGAAGGGAACTGGAGAACTCAGCCTTCATCTTCCCCCA 3246

Qy   1576 cccccatttgctggttcagctatgattcaggtaggctctgccctgggccatgacaccagc 1635
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3245 CCCCCATTTGCTGGTTCAGCTATGATTCAGGTAGGCTCTGCCCTGGGCCATGACACCAGC 3186

Qy   1636 ctcttagtcttc 1647
          ||||||||||||
Db   3185 CTCTTAGTCTTC 3174
```

Search completed: October 3, 2002, 16:30:41
Job time: 14311 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 16:14:49 ; Search time 5701.1 Seconds
(without alignments)
3899.157 Million cell updates/sec

Title: US-09-826-581-5
Perfect score: 1647
Sequence: 1 ttggtctggggctggccaca.....acaccagcctcttagtcttc 1647

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	456.4	27.7	572	10	BI344527	BI344527 373008 MA
2	384	23.3	970	9	AL533446	AL533446 AL533446
3	380	23.1	758	9	AJ396118	AJ396118 AJ396118
4	374.6	22.7	685	9	BB630381	BB630381 BB630381
5	366.8	22.3	921	9	AL548987	AL548987 AL548987
6	363.2	22.1	905	9	AL555228	AL555228 AL555228
7	361	21.9	940	9	AL554278	AL554278 AL554278
8	354	21.5	548	10	BI775360	BI775360 467815 MA
9	350	21.3	908	9	AL552459	AL552459 AL552459
10	346.8	21.1	864	10	BI858240	BI858240 603384001
11	343	20.8	826	9	AL519198	AL519198 AL519198
12	342.8	20.8	951	9	AL513784	AL513784 AL513784
13	336.6	20.4	1013	10	BM462694	BM462694 AGENCOURT
14	335	20.3	1004	10	BM470188	BM470188 AGENCOURT
15	329.2	20.0	595	10	BM487789	BM487789 pgm2n.pk0
16	328.8	20.0	967	9	AL524822	AL524822 AL524822
17	328.6	20.0	1006	10	BM469833	BM469833 AGENCOURT
18	316.4	19.2	870	10	BI910928	BI910928 603069326
19	312.8	19.0	647	10	BM440762	BM440762 pgrln.pk0
20	310.2	18.8	784	9	AU080009	AU080009 AU080009
21	308.4	18.7	841	10	BI546514	BI546514 603191524
22	303.4	18.4	633	10	BJ072114	BJ072114 BJ072114
23	303.4	18.4	808	10	BI771184	BI771184 603054888
24	301	18.3	636	10	BG713637	BG713637 pgl1n.pk0
25	298	18.1	782	10	BI819312	BI819312 603037761
26	295.6	17.9	444	10	BF890374	BF890374 291826 MA
27	293.4	17.8	761	9	AL632024	AL632024 AL632024
28	287.6	17.5	717	10	BI152149	BI152149 602915522
29	287.6	17.5	849	10	BI657718	BI657718 603283104
30	286.2	17.4	795	10	BE871189	BE871189 601448767
31	285.8	17.4	413	9	AA178898	AA178898 zp38d10.r
32	285.2	17.3	655	9	BB629521	BB629521 BB629521
33	285	17.3	934	10	BI256222	BI256222 602975031
34	283	17.2	651	10	BJ072368	BJ072368 BJ072368
35	282.4	17.1	782	10	BI821538	BI821538 603038457
36	282	17.1	903	10	BG177822	BG177822 602314222
37	280.2	17.0	422	9	AW427435	AW427435 63185 MAR
38	280	17.0	872	10	BG821178	BG821178 602724586
39	276.2	16.8	708	9	AW411228	AW411228 fh11b02.x
40	275.8	16.7	928	10	BI091156	BI091156 602854771
41	274.2	16.6	684	10	BG101440	BG101440 uy72c12.y
42	273.8	16.6	694	10	BF468462	BF468462 UI-M-CD0-
43	273.2	16.6	710	10	BI223706	BI223706 602943249
44	272.6	16.6	616	9	BB626865	BB626865 BB626865
45	270	16.4	399	9	AW356079	AW356079 38073 MAR

ALIGNMENTS

RESULT 1
BI344527
LOCUS BI344527 572 bp mRNA linear EST 30-JUL-2001
DEFINITION 373008 MARC 2PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BI344527
VERSION BI344527.1 GI:15037807
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 572)
AUTHORS Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W. and Keele,J.W.
TITLE Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR PRimers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCCAGTCACGACG
Plate: 119 row: I column: 11
Seq primer: ATTTAGGTGACACTATAG.
FEATURES Location/Qualifiers
source 1. .572
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2PIG"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;

Library made from pooled tissue from testis, ovary, endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT 106 a 190 c 151 g 124 t 1 others
ORIGIN

Query Match 27.7%; Score 456.4; DB 10; Length 572;
Best Local Similarity 89.8%; Pred. No. 4.6e-97;
Matches 501; Conservative 0; Mismatches 56; Indels 1; Gaps 1;

```
Qy    853 aggctgcttcaagcctctggtctccatctctcctaatgatagcctgtttgaagctgtcta 912
          ||||||||||||||||||||||||||||||||||||||| ||||| |||||||| ||||||||||
Db     15 AGGCTGCTTCAAGCCTCTGGTCTCCATCTCTCCCAATGACAGCCTGTTCGAAGCTGTCTA 74

Qy    913 caccctcatcaagaaccggatccatcgcctgcctgttcttgacccggtgtcaggcaacgt 972
          | |||||||||||||||||||||| |||||||| || || ||||| || || ||    ||
Db     75 CGCCCTCATCAAGAACCGGATCCACCGCCTGCCGGTCCTGGACCCTGTCTCCGGGGCTGT 134

Qy    973 actccacatcctcacacacaaacgcctgctcaagttcctgcacatctttggttccctgct 1032
          |||||||||||||||| || || || ||||||||||||||||||||||  |||||||
Db    135 GCTCCACATCCTCACACATAAGCGGCTTCTCAAGTTCCTGCACATCTTTGGCACCCTGCT 194

Qy   1033 gccccggccctccttcctctaccgcactatccaagatttgggcatcggcacattccgaga 1092
          ||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    195 GCCCCGGCCCTCCTTCCTCTACCGCACCATCCAAGATTTGGGCATCGGCACATTCCGAGA 254

Qy   1093 cttggctgtggtgctggagacagcacccatcctgactgcactggacatctttgtggaccg 1152
          |||||| |||||||||||| || || |||||||||||  ||||||||||||| ||||||||
Db    255 CTTGGCCGTGGTGCTGGAAACGGCGCCCATCCTGACCGCACTGGACATCTTCGTGGACCG 314

Qy   1153 gcgtgtgtctgcactgcctgtggtcaacgaatgtggtcaggtcgtgggcctctattcccg 1212
          |||||||||||| |||||||||||||||||  ||| ||||| |||||||||||| || ||
Db    315 GCGTGTGTCTGCGCTGCCTGTGGTCAACGAAACTGGACAGGTAGTGGGCCTCTACTCTCG 374

Qy   1213 ctttgatgtgattcacctggctgcccagcaaacctacaaccacctggacatgagtgtggg 1272
          |||||||||||| |||||||||||||||| ||||| ||||||||||||||||||||||||||
Db    375 CTTTGATGTGATCCACCTGGCTGCCCAACAAACATACAACCACCTGGACATGAGTGTGGG 434

Qy   1273 agaagccctgaggcagaggacactatgtctggagggagtcctttcctgccagccccacga 1332
          |||||||||||||||| ||||||| |||||||| || ||||||||||||||||||||||||
Db    435 AGAAGCCCTGAGGCAGCGGACACTGTGTCTGGAAGGCGTCCTTTCCTGCCAGCCCCACGA 494

Qy   1333 gagcttgg-gggaagtgatcgacaggattgctcgggagcaggtacacaggctggtgctag 1391
          || ||||| ||||||| || ||| ||||||  ||||| |||| ||| | |||||||| |
Db    495 GACCTTGGNGGGAAGTCATTGACCGGATTGTCCGGGAACAGGTGCACCGCCTGGTGCTCG 554

Qy   1392 tggacgagacccagcatc 1409
          |||| |||||||||||| |
Db    555 TGGATGAGACCCAGCACC 572
```

RESULT 2
AL533446
LOCUS AL533446 970 bp mRNA linear EST 13-FEB-2001
DEFINITION AL533446 LTI_FL015_Brn1 Homo sapiens cDNA clone CS0DN004YN03 5 prime, mRNA sequence.
ACCESSION AL533446
VERSION AL533446.1 GI:12796939
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 970)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES Location/Qualifiers
source 1. .970
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DN004YN03"
/clone_lib="LTI_FL015_Brn1"
/sex="male"
/tissue_type="Adult brain"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT 255 a 230 c 217 g 268 t
ORIGIN

Query Match 23.3%; Score 384; DB 9; Length 970;
Best Local Similarity 64.8%; Pred. No. 6.8e-80;
Matches 570; Conservative 0; Mismatches 310; Indels 0; Gaps 0;

```
Qy    567 tctacatgcgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagc 626
          | || |    |||||| ||   ||   |||||| ||    ||  | || ||||||||
Db     90 TGTATACTTCCTTCATGAAGTCTCATCGCTGCTATGACCTGATTCCCACAAGCTCCAAAT 149

Qy    627 tagtcatcttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacg 686
          | ||  | || || ||    ||| || | ||||| || || |||||| ||||| | ||||
Db    150 TGGTTGTATTTGATACGTCCCTGCAGGTGAAGAAAGCTTTTTTTGCTTTGGTGACTAACG 209

Qy    687 gtgtgcgggcagcccctctatgggacagcaagaagcagagctttgtggggatgctgacca 746
          |||| || || |||||| |||||||| || |||||||| || |||||||| ||||||||||
Db    210 GTGTACGAGCTGCCCCTTTATGGGATAGTAAGAAGCAAAGTTTTGTGGGCATGCTGACCA 269

Qy    747 tcactgacttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatg 806
          ||||||| ||||||    | ||||| |||||||| |  ||  || |||| ||||||||||
Db    270 TCACTGATTTCATCAATATCCTGCACCGCTACTATAAATCAGCCTTGGTACAGATCTATG 329

Qy    807 agattgaacaacataagattgagacctggagggagatctacctgcaaggctgcttcaagc 866
          || | ||| |||| ||||| || || ||||| ||| | || || || | || ||| || |
Db    330 AGCTAGAAGAACACAAGATAGAAACTTGGAGAGAGGTGTATCTCCAGGACTCCTTTAAAC 389

Qy    867 ctctggtctccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaaga 926
          | || |||| ||| ||||||||||  ||| ||||||| |||||||   |  | ||   ||
Db    390 CGCTTGTCTGCATTTCTCCTAATGCCAGCTTGTTTGATGCTGTCTCTTCATTAATTCGGA 449

Qy    927 accggatccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctca 986
          ||  ||||||  | ||||| ||| ||||||| |  ||||||||     |  |||||||||
Db    450 ACAAGATCCACAGGCTGCCAGTTATTGACCCAGAATCAGGCAATACTTTGTACATCCTCA 509

Qy    987 cacacaaacgcctgctcaagttcctgcacatctttggttccctgctgccccggcctcct 1046
          | ||||| ||| | || ||||||||  |  | |||     |   | | |||  |||    |
Db    510 CCCACAAGCGCATTCTGAAGTTCCTCAAATTGTTTATCACTGAGTTCCCCAAGCCAGAGT 569

Qy   1047 tcctctaccgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgc 1106
          || | | |    || |  |||| |    || ||||| |       |  | ||| |||| |
Db    570 TCATGTCCAAGTCTCTGGAAGAGCTACAGATTGGCACCTATGCCAATATTGCTATGGTTC 629

Qy   1107 tggagacagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcac 1166
               ||  | ||| ||      || |||| || |||||  | |  || || || || |
Db    630 GCACTACCACCCCCGTCTATGTGGCTCTGGGGATTTTTGTACAGCATCGAGTCTCAGCCC 689

Qy   1167 tgcctgtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattc 1226
          |||| |||||  | ||   || |  || |||| | |||| |||   |||||||| ||
Db    690 TGCCAGTGGTGGATGAGAAGGGGCGTGTGGTGGACATCTACTCCAAGTTTGATGTTATCA 749

Qy   1227 acctggctgcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggc 1286
          | ||||| ||  |  | |||||||||| |||| ||  |   ||||    ||||| ||   |
Db    750 ATCTGGCAGCAGAAAAGACCTACAACAACCTAGATGTATCTGTGACTAAAGCCTTGCAAC 809

Qy   1287 agaggacactatgtctggagggagtcctttcctgccagccccacgagagcttgggggaag 1346
```

```
Db    810 ATCGATCACATTACTTTGAGGGTGTTCTCAAGTGCTACCTGCATGAGACTCTGGAGACCA 869

Qy   1347 tgatcgacaggattgctcgggagcaggtacacaggctggtgctagtggacgagacccagc 1406
Db    870 TCATCAACAGGCTAGTGGAAGCAGAGGTTCACCGACTTGTAGTGGTGGATGAAAATGATG 929

Qy   1407 atctcttgggcgtggtctccctctccgacatccttcaggc 1446
Db    930 TGGTCAAGGGAATTGTATCACTGTCTGACATCCTGCAGGC 969


RESULT  3
AJ396118
LOCUS       AJ396118                758 bp    mRNA    linear   EST 25-JAN-2001
DEFINITION  AJ396118 dkfz426 Gallus gallus cDNA clone 25f16r1, mRNA sequence.
ACCESSION   AJ396118
VERSION     AJ396118.1  GI:7127728
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 758)
  AUTHORS   Abdrakhmanov,I., Lodygin,D., Geroth,P., Arakawa,H., Law,A., Plachy
            ,J., Korn,B. and Buerstedde,J.M.
  TITLE     A large database of chicken bursal ESTs as a resource for the
            analysis of vertebrate gene function
  JOURNAL   Genome Res. 10 (12), 2062-2069 (2000)
  MEDLINE   20568495
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinistr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES             Location/Qualifiers
     source          1. .758
                     /organism="Gallus gallus"
                     /strain="CB"
                     /db_xref="taxon:9031"
                     /clone="25f16r1"
                     /clone_lib="dkfz426"
                     /tissue_type="Bursa of Fabricius"
BASE COUNT      168 a    238 c    186 g    166 t
ORIGIN


  Query Match           23.1%;  Score 380;  DB 9;  Length 758;
  Best Local Similarity   69.3%;  Pred. No. 5.3e-79;
  Matches  518;  Conservative    0;  Mismatches  230;  Indels    0;  Gaps    0;

Qy    614 actagctccaagctagtcatcttcgacaccatgctggagatcaagaaggccttctttgct 673
Db      1 ACCAGCTCCAAACTCGTCGTCTTCGACACTTCCCTTCAGGTGAAGAAGGCTTTCTTTGCA 60

Qy    674 ctggtggccaacggtgtgcgggcagcccctctatgggacagcaagaagcagagctttgtg 733
Db     61 CTGGTCACCAATGGCGTGCGGGCAGCCCCGCTGTGGGACAGCAAAAAGCAGAGCTTTGTG 120

Qy    734 gggatgctgaccatcactgacttcatcctggtgctgcatcgctactacaggtccccccctg 793
Db    121 GGCATGCTGACCATCACTGACTTCATCAACATCCTGCACCGCTACTACAAGTCGCCCATG 180

Qy    794 gtccagatctatgagattgaacaacataagattgagacctggagggagatctacctgcaa 853
Db    181 GTGCAGATCTATGAGCTGGAGGAGCACAAAATCGAGACGTGGAGAGAGGTGTACCTGCAG 240

Qy    854 ggctgcttcaagcctctggtctccatctctcctaatgatagcctgtttgaagctgtctac 913
Db    241 GACTCCTTTAAGCCGTTGGTCTGCATCTCCCCCAATGCCAGCCTTTTTGATGCCGTCTCC 300

Qy    914 accctcatcaagaaccggatccatcgcctgcctgttcttgacccggtgtcaggcaacgta 973

Db    301 TCCCTGATCCGCAATAAGATCCACCGCCTCCCCGTCATCGACCCCGACTCGGGCAACACT 360

Qy    974 ctccacatcctcacacacaaacgcctgctcaagttcctgcacatctttggttccctgctg 1033
Db    361 CTCTACATCCTCACCCCCAAACGCATCCTCAAGTTCCTCAAACTCTTTATTGCAGAGGTC 420

Qy   1034 ccccggccctccttcctctaccgcactatccaagatttgggcatcggcacattccgagac 1093
Db    421 CCAAAGCCTGAGTTTATGGCCCGGACTTTGGAGGAGCTGCAGATCGGTACCTACAGCAAC 480

Qy   1094 ttggctgtggtgctggagacagcacccatcctgactgcactggacatctttgtggaccgg 1153
Db    481 ATCGCCGTGGTGAGCACCAGCACTCCCATCTATGTGGCTCTGGGCATCTTCGTGCAGCAC 540

Qy   1154 cgtgtgtctgcactgcctgtggtcaacgaatgtggtcaggtcgtgggcctctattcccgc 1213
Db    541 CGCGTCTCCGCTCTGCCCGTGGTCGATGATTCGGGGCGGGTGGTGGATATCTACTCCAAG 600

Qy   1214 tttgatgtgattcacctggctgcccagcaaacctacaaccacctggacatgagtgtggga 1273
Db    601 TTCGACGTTATCAATTTGGCAGCTGAGAAGACCTACAACAACCTGGACGTGACGGTGACG 660

Qy   1274 gaagccctgaggcagaggacactatgtctggagggagtcctttcctgccagccccacgag 1333
Db    661 CGGGCGCTGCAGCACCGCTCCCATTACTTTGAGGGCGTCCTCAAATGCTACAAACACGAA 720

Qy   1334 agcttgggggaagtgatcgacaggattg 1361
Db    721 ACCTTGGAAGCCATCATCAACCGCCTGG 748


RESULT  4
BB630381
LOCUS       BB630381                685 bp    mRNA    linear   EST 26-OCT-2001
DEFINITION  BB630381 RIKEN full-length enriched, 6 days neonate skin Mus
            musculus cDNA clone A030014A04 5', mRNA sequence.
ACCESSION   BB630381
VERSION     BB630381.1  GI:16467364
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 685)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
```

sequencing pipeline with 384 multicapillary sequencer. Genome Res. . 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara ,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
e mouse tissues.

```
FEATURES             Location/Qualifiers
     source          1. .685
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="A030014A04"
                     /clone_lib="RIKEN full-length enriched, 6 days neonate
                     skin"
                     /tissue_type="skin"
                     /dev_stage="6 days neonate"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 10.0 and subtraction to Rot = 185.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC
                     3']. cDNA was cleaved with XhoI and BamHI. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I. "
BASE COUNT      165 a    199 c    198 g    123 t
ORIGIN
```

```
  Query Match             22.7%;  Score 374.6;  DB 9;  Length 685;
  Best Local Similarity   77.7%;  Pred. No. 9.4e-78;
  Matches  492;  Conservative    0;  Mismatches  134;  Indels    7;  Gaps    3;

Qy       8 ggggctggccacatggagcccgggctggagcacgcactgcgcaggacccctttcctggagc 67
           || ||||||| |||||||||||| |||||||||| || |||   ||||||   ||||||||
Db       1 GGAGCTGGCCGCATGGAGCCCGAGCTGGAGCACACATTGCCTGGGACCCTGACCTGGAGC 60

Qy      68 agccttgggggttctgagcatcaagagatgagcttcctagagcaagaaaacagcagctca 127
             |  |||||||  | |||  ||||||||||  |||| |||| |||| | |    | ||||
Db      61 CACAGTGGGGGTCCAGAGAGTCAAGAGATGGACTTCTTAGAACAAGGAGA---AAACTCA 117

Qy     128 tggccatcaccagctgtgaccagcagctcagaaagaatccgtgggaaacggagggccaaa 187
           ||||| |||||||||||| ||| |||||||||||||| | |||  | |||| | | |||
Db     118 TGGCCCTCACCAGCTGTGGCCACCAGCTCAGAAAGAACCTGTGCCATACGGGGAGTCAAG 177

Qy     188 gccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtcaggggaa 247
           || |  |||||||| || ||| || | || |||||||  || |||||||||  ||| |||
Db     178 GCTTCCAGATGGACGAGACAGGAGGCCGTAGAGGAAGCAGAACCACCAGGTTTGGGAGAA 237

Qy     248 ggtccccggtccaggccagctgctgagtccaccgggctggaggccacattccccaagacc 307
           ||| ||| |||||  ||||||||||||||||||| ||| ||||||||||||||| ||| ||
Db     238 GGTGCCCAGTCCATACCAGCTGCTGAGTCCACCAGGCAGGAGGCCACATTCCCGAAGGCC 297

Qy     308 acacccttggctcaagctgatcct---gccggggtgggcactccaccaacagggtgggac 364
           |||||||||||||||||| |||    || |  | ||  ||  | || ||||||||||
Db     298 ACACCCTTGGCTCAAGCTGTTCCCTTGGCTGAAGCGGAGACCTCCCCCACAGGGTGGGAC 357

Qy     365 tgcctcccctctgactgtacagcctcagctgcaggctccagcacagatgatgtggagctg 424
             |||    | |||||| |||||||||| |  |||||||||||  |||  ||||||||
Db     358 CTGCTCTTGCCCGACTGTGCAGCCTCAGCAGGGGGCTCCAGCACAAGTGACCTGGAGCTG 417

Qy     425 gccacggagttcccagccacagaggcctgggagtgtgagctagaaggcctgctggaagag 484
             |||  |||||||||||| |||||||||||| |||||||| |||||||||  | ||||
Db     418 ACCATAGAGTTCCCAGCCCCAGAGGCCTGGGACTGTGAGCTGGAAGGCCTGGGGAAAGAC 477

Qy     485 aggcctgccctgtgcctgtccccgcaggccccatttcccaagctgggctgggatgacgaa 544
           ||||||   |   | |  ||||| |||||||||| |||     ||| | |||||||||| ||
Db     478 AGGCCTCGTCCTGGTCCATCCCCACAGGCCCCACTTCTTGGCCTGAGTTGGGATGACCAA 537

Qy     545 ctgcggaaacccggcgcccagatctacatgcgcttcatgcaggagcacacctgctacgat 604
           || |  |||||||| |||||| |||||||||| ||| |||||||| |||||||| || |||
Db     538 CTTCACAAACCCGGAGCCCAGGTCTACATGCACTTTATGCAGGAACACACCTGTTATGAT 597

Qy     605 gccatggc-aactagctccaagctagtcatctt 636
           ||||||||  || ||||||||  | ||||||||
Db     598 GCCATGGCTTACCAGCTCCAAATTGGTCATCTT 630
```

```
RESULT   5
AL548987
LOCUS       AL548987                 921 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL548987 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI042YD16 5
            prime, mRNA sequence.
ACCESSION   AL548987
VERSION     AL548987.1  GI:12884536
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 921)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .921
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DI042YD16"
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer.  Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      243 a    217 c    206 g    255 t
ORIGIN
```

```
  Query Match             22.3%;  Score 366.8;  DB 9;  Length 921;
  Best Local Similarity   65.0%;  Pred. No. 7.5e-76;
  Matches  542;  Conservative    0;  Mismatches  292;  Indels    0;  Gaps    0;

Qy     567 tctacatgcgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagc 626
           | || |    ||||||| ||    ||   |||||| ||    ||  | || ||||||||
Db      67 TGTATACTTCCTTCATGAAGTCTCATCGCTGCTATGACCTGATTCCCACAAGCTCCAAAT 126

Qy     627 tagtcatcttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacg 686
           | ||  | || || ||    ||| || | ||||| || || |||||| ||||| | ||||
```

```
Db    127 TGGTTGTATTTGATACGTCCCTGCAGGTGAAGAAAGCTTTTTTTGCTTTGGTGACTAACG 186

Qy    687 gtgtgcgggcagcccctctatgggacagcaagaagcagagctttgtggggatgctgacca 746
          |||| || || |||||| |||||||| || |||||||| || |||||||| ||||||||||
Db    187 GTGTACGAGCTGCCCCTTTATGGGATAGTAAGAAGCAAAGTTTTGTGGGCATGCTGACCA 246

Qy    747 tcactgacttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatg 806
          ||||||| ||||||    | ||||| |||||||| |   ||  || |||| ||||||||||
Db    247 TCACTGATTTCATCAATATCCTGCACCGCTACTATAAATCAGCCTTGGTACAGATCTATG 306

Qy    807 agattgaacaacataagattgagacctggagggagatctacctgcaaggctgcttcaagc 866
          || | ||| |||| ||||| || || ||||| ||| | || || || | || ||| || |
Db    307 AGCTAGAAGAACACAAGATAGAAACTTGGAGACAGGTGTATCTCCAGGACTCCTTTAAAC 366

Qy    867 ctctggtctccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaaga 926
          | || |||| ||| |||||||||| ||| ||||||| |||||||   |  | ||   ||
Db    367 CGCTTGTCTGCATTTCTCCTAATGCCAGCTTGTTTGATGCTGTCTCTTCATTAATTCGGA 426

Qy    927 accggatccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctca 986
          ||  ||||||  | ||||| ||| ||||||| |  ||||||||     |  |||||||||
Db    427 ACAAGATCCACAGGCTGCCAGTTATTGACCCAGAATCAGGCAATACTTTGTACATCCTCA 486

Qy    987 cacacaaacgcctgctcaagttcctgcacatctttggttccctgctgccccggccctcct 1046
          | ||||| ||| | || ||||||||  |  | |||    |   | | ||| |||    |
Db    487 CCCACAAGCGCATTCTGAAGTTCCTCAAATTGTTTATCACTGAGTTCCCCAAGCCAGAGT 546

Qy   1047 tcctctaccgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgc 1106
          || | | |    || |  |||| |    || ||||| |       |  | ||| |||| |
Db    547 TCATGTCCAAGTCTCTGGAAGAGCTACAGATTGGCACCTATGCCAATATTGCTATGGTTC 606

Qy   1107 tggagacagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcac 1166
               ||  | ||| ||       || |||| || |||||  | |  || || || || |
Db    607 GCACTACCACCCCCGTCTATGTGGCTCTGGGGATTTTTGTACAGCATCGAGTCTCAGCCC 666

Qy   1167 tgcctgtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattc 1226
          |||| |||||  | ||    || |  || |||| | |||| |||   |||||||| ||
Db    667 TGCCAGTGGTGGATGAGAAGGGGCGTGTGGTGGACATCTACTCCAAGTTTGATGTTATCA 726

Qy   1227 acctggctgcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggc 1286
          | ||||| ||  |  | |||||||||| |||| ||  |  ||||   ||||| ||   |
Db    727 ATCTGGCAGCAGAAAAGACCTACAACAACCTAGATGTATCTGTGACTAAAGCCTTGCAAC 786

Qy   1287 agaggacactatgtctggagggagtcctttcctgccagccccacgagagcttgggggaag 1346
          |  |  |||  |   | ||||| || ||    ||| | |  || ||||  ||| |
Db    787 ATCGATCACATTACTTTGAGGGTGTTCTCAAGTGCTACCTGCATGAGACTCTGGAGACCA 846

Qy   1347 tgatcgacaggattgctcgggagcaggtacacaggctggtgctagtggacgaga 1400
          | ||| ||||| | |      |   |||| ||| | || ||  | ||||| || |
Db    847 TCATCAACAGGCTAGTGGAAGCAGAGGTTCACCGACTTGTAGTGGTGGATGAAA 900


RESULT   6
AL555228
LOCUS       AL555228                905 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL555228 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DK007YE02 5
            prime, mRNA sequence.
ACCESSION   AL555228
VERSION     AL555228.1  GI:12896759
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 905)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .905
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DK007YE02"
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer.  Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      237 a    218 c    199 g    250 t      1 others
ORIGIN


  Query Match             22.1%;  Score 363.2;  DB 9;  Length 905;
  Best Local Similarity   64.9%;  Pred. No. 5.2e-75;
  Matches  536;  Conservative    1;  Mismatches  289;  Indels    0;  Gaps    0;

Qy    567 tctacatgcgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagc 626
          | || |    |||||||| ||   ||   |||||| ||    ||  | || ||||||||
Db     80 TGTATACTTCCTTCATGAAGTCTCATCGCTGCTATGACCTGATTCCCACAAGCTCCAAAT 139

Qy    627 tagtcatcttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacg 686
          | ||   | || || ||     ||| || | ||||| || || |||||| ||||| | ||||
Db    140 TGGTTGTATTTGATACGTCCCTGCAGGTGAAGAAAGCTTTTTTTGCTTTGGTGACTAACG 199

Qy    687 gtgtgcgggcagcccctctatgggacagcaagaagcagagctttgtggggatgctgacca 746
          |||| || || |||||| ||||||| || |||||||| || |||||||| ||||||||||
Db    200 GTGTACGAGCTGCCCCTTTATGGGATAGTAAGAAGCAAAGTTTTGTGGGCATGCTGACCA 259

Qy    747 tcactgacttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatg 806
          |||||||  ||||||    | ||||| |||||||| |   ||  || |||| ||||||||||
Db    260 TCACTGATTTCATCAATATCCTGCACCGCTACTATAAATCAGCCTTGGTACAGATCTATG 319

Qy    807 agattgaacaacataagattgagacctggagggagatctacctgcaaggctgcttcaagc 866
          || | ||| |||| ||||| || || ||||| ||| | || || || | || ||| || |
Db    320 AGCTAGAAGAACACAAGATAGAAACTTGGAGAGAGGTGTATCTCCAGGACTCCTTTAAAC 379

Qy    867 ctctggtctccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaaga 926
          | || |||| ||| |||||||||| ||| ||||||| |||||||   |  | ||   ||
Db    380 CGCTTGTCTGCATTTCTCCTAATGCCAGCTTGTTTGATGCTGTCTCTTCATTAATTCGGA 439

Qy    927 accggatccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctca 986
          ||  ||||||  | ||||| ||| ||||||| |  ||||||||     |  |||||||||
Db    440 ACAAGATCCACAGGCTGCCAGTTATTGACCCAGAATCAGGCAATACTTTGTACATCCTCA 499

Qy    987 cacacaaacgcctgctcaagttcctgcacatctttggttccctgctgccccggccctcct 1046
          | ||||| ||| | || ||||||||  |  | |||    |   | | ||| |||    |
Db    500 CCCACAAGCGCATTCTGAAGTTCCTCAAATTGTTTATCACTGAGTTCCCCAAGCCAGAGT 559

Qy   1047 tcctctaccgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgc 1106
          || | | |    || |  |||| |    || ||||| |       |  | ||| |||| |
Db    560 TCATGTCCAAGTCTCTGGAAGAGCTACAGATTGGCACCTATGCCAATATTGCTATGGTTC 619

Qy   1107 tggagacagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcac 1166
               ||  | ||| ||       || |||| || |||||  | |  || || || || |
Db    620 GCACTACCACCCCCGTCTATGTGGCTCTGGGGATTTTTGTACAGCATCGAGTCTCAGCCC 679

Qy   1167 tgcctgtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattc 1226
          |||| |||||  | ||    || |  || |||| | |||| |||   |||||||| ||
Db    680 TGCCAGTGGTGGATGAGAAGGGGCGTGTGGTGGACATCTACTCCAAGTTTGATGTTATCA 739

Qy   1227 acctggctgcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggc 1286
          | ||||| ||  |  | |||||||||| |||| ||  |  ||||   ||||| ||   |
Db    740 ATCTGGCAGCAGAAAAGACCTACAACAACCTAGATGTATCTGTGACTAAAGCCTTGCAAC 799

Qy   1287 agaggacactatgtctggagggagtcctttcctgccagccccacgagagcttgggggaag 1346
          |  |  |||  |   | ||||| || ||    ||| | |  || ||||  ||| |
Db    800 ATCGATCACATTACTTTGAGGGTGTTCTCAAGTGCTACCTGCATGAGACTCTGGAGACCA 859
```

```
Db    800 ATCGATCACATTACTTTGAGGGTGTTCTCAAGTGCTACCTGCATGAGACTCTGGAGACCA 859

Qy   1347 tgatcgacaggattgctcgggagcaggtacacaggctggtgctagt 1392
          | ||| ||||| | |   : |   |||| ||| | || ||  | ||
Db    860 TCATCAACAGGCTAGTGGRAGCAGAGGTTCACCGACTTGTAGTGGT 905


RESULT   7
AI554278
LOCUS       AL554278                940 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL554278 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI082YL23 5
            prime, mRNA sequence.
ACCESSION   AL554278
VERSION     AL554278.1  GI:12894901
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 940)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .940
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DI082YL23"
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer.  Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      248 a    222 c    208 g    260 t      2 others
ORIGIN


  Query Match             21.9%;  Score 361;  DB 9;  Length 940;
  Best Local Similarity   64.8%;  Pred. No. 1.8e-74;
  Matches  535;  Conservative    0;  Mismatches  290;  Indels    0;  Gaps    0;

Qy    567 tctacatgcgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagc 626
          | || |    ||||||| ||   ||   |||||| ||    ||  | || ||||||||
Db     93 TGTATACTTCCTTCATGAAGTCTCATCGCTGCTATGACCTGATTCCCACAAGCTCCAAAT 152

Qy    627 tagtcatcttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacg 686
          | ||  | || || ||    ||| || | ||||| || || |||||| ||||| | ||||
Db    153 TGGTTGTATTTGATACGTCCCTGCAGGTGAAGAAAGCTTTTTTTGCTTTGGTGACTAACG 212

Qy    687 gtgtgcgggcagcccctctatgggacagcaagaagcagagctttgtggggatgctgacca 746
          |||| || || |||||| ||||||| || |||||||| || |||||||| ||||||||||
Db    213 GTGTACGAGCTGCCCCTTTATGGGATAGTAAGAAGCAAAGTTTTGTGGGCATGCTGACCA 272

Qy    747 tcactgacttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatg 806
          ||||||| ||||||    | ||||| |||||||| |  ||  || |||| ||||||||||
Db    273 TCACTGATTTCATCAATATCCTGCACCGCTACTATAAATCAGCCTTGGTACAGATCTATG 332

Qy    807 agattgaacaacataagattgagacctggagggagatctacctgcaaggctgcttcaagc 866
          || | ||| |||| ||||| || || ||||| ||| | || || || | || ||| || |
Db    333 AGCTAGAAGAACACAAGATAGAAACTTGGAGAGAGGTGTATCTCCAGGACTCCTTTAAAC 392

Qy    867 ctctggtctccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaaga 926
          | || |||| ||| |||||||||  ||| ||||||| |||||||   |  | ||   ||
Db    393 CGCTTGTCTGCATTTCTCCTAATGCCAGCTTGTTTGATGCTGTCTCTTCATTAATTCGGA 452

Qy    927 accggatccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctca 986
          ||  ||||||  | ||||| ||| ||||||| |  ||||||||     |  |||||||||
Db    453 ACAAGATCCACAGGCTGCCAGTTATTGACCCAGAATCAGGCAATACTTTGTACATCCTCA 512

Qy    987 cacacaaacgcctgctcaagttcctgcacatctttggttccctgctgccccggccctcct 1046
          | ||||| ||| | || ||||||||  |  | |||    |   | | |||  |||    |
Db    513 CCCACAAGCGCATTCTGAAGTTCCTCAAATTGTTTATCACTGAGTTCCCCAAGCCAGAGT 572

Qy   1047 tcctctaccgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgc 1106
          || | | |    || |  ||||  |    || ||||| |       |  | ||| |||| |
Db    573 TCATGTCCAAGTCTCTGGAAGAGCTACAGATTGGCACCTATGCCAATATTGCTATGGTTC 632

Qy   1107 tggagacagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcac 1166
               ||  | ||| ||       || ||||  || |||||  | |  || || || || |
Db    633 GCACTACCACCCCCGTCTATGTGGCTCTGGGGATTTTTGTACAGCATCGAGTCTCAGCCC 692

Qy   1167 tgcctgtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattc 1226
          |||| |||||  | ||    || |  || |||| | |||| |||   |||||||| ||
Db    693 TGCCAGTGGTGGATGAGAAGGGGCGTGTGGTGGACATCTACTCCAAGTTTGATGTTATCA 752

Qy   1227 acctggctgcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggc 1286
          | ||||| ||  |  | ||||||||| |||| ||  |  ||||    ||||| ||   |
Db    753 ATCTGGCAGCAGAAAAGACCTACAACAACCTAGATGTATCTGTGACTAAAGCCTTGCAAC 812

Qy   1287 agaggacactatgtctggagggagtcctttcctgccagccccacgagagcttgggggaag 1346
          |  |  |||  |   | |||||| || ||    ||| | |  || ||||   ||| |
Db    813 ATCGATCACATTACTTTGAGGGTGTTCTCAAGTGCTACCTGCATGAGACTCTGGAGACCA 872

Qy   1347 tgatcgacaggattgctcgggagcaggtacacaggctggtgctag 1391
          | ||| ||||| | |     |   |||| ||| | || ||  | |
Db    873 TCATCAACAGGCTAGTGGAAGCAGAGGTTCACCGACTTGTAGTKG 917


RESULT   8
BI775360
LOCUS       BI775360                548 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  467815 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BI775360
VERSION     BI775360.1  GI:15776346
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 548)
  AUTHORS   Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
            ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
            Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
            Keele,J.W.
  TITLE     Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
  JOURNAL   Genome Res. 11 (4), 626-630 (2001)
  MEDLINE   21180013
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 90  row: I  column: 16
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
```

```
     source          1. .548
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="MARC 2BOV"
                     /tissue_type="pooled"
                     /lab_host="DH10H"
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled tissue from testis, thymus,
                     semitendonosus muscle, longissimus muscle, pancreas,
                     adrenal, and endometrium."
BASE COUNT      119 a    173 c    163 g     93 t
ORIGIN


  Query Match             21.5%;  Score 354;  DB 10;  Length 548;
  Best Local Similarity   81.5%;  Pred. No. 6.1e-73;
  Matches  448;  Conservative    0;  Mismatches   95;  Indels    7;  Gaps    3;

Qy     166 ccgtgggaaacggagggccaaagccttgag-atggacaaggcagaagtcggtggaggaag 224
           || |||| ||| |||| |||| | ||  || ||||| ||| ||| ||   ||  |||||
Db       2 CCTTGGGGAACAGAGGACCAAGGTCTCAAGTATGGAAAAGCCAGGAGGATGTAAAGGAAA 61

Qy     225 gggagccaccaggtcaggggggaaggtccccggtccaggccagctgctgagtccaccgggc 284
           |||||| |||| | || || |||||||| |||| | ||||||||||||||||||||||
Db      62 GGGAGCTGCCAGGCCTGGAGGGAGGTCCCCATTCCAAGGCAGCTGCTGAGTCCACCGGGC 121

Qy     285 tggaggccacattccccaagaccacaccctggctcaagctgatcc---tgccggggtgg 341
           |||||||||||||||||| |||||||||||| |||||   |||     ||| |||||
Db     122 TGGAGGCCACATTCCCCAAGGCCACACCCTTGGCCCAAGCCACTCCCTTGTCCGCGGTGG 181

Qy     342 gcactccaccaacagggtgggactgcctcccctctgactgtacagcctcagctgcaggct 401
           |||| ||  |||||   | ||| |||||||||||||||||||||||| ||| |  |||
Db     182 GCACCCCCACAACAGAACGAGACAGCCTCCCCTCTGACTGTACAGCCTCCGCTTCCAGCT 241

Qy     402 ccagcacagatgatgtggagctggccacggagttcccagccacagaggcctgggagtgtg 461
           |||||||||| ||| |||| | || ||   ||||| ||||| |||  || |||| |  ||
Db     242 CCAGCACAGACGATCTGGATCAGGGCATATAGTTCTCAGCCCCAGCAGCGTGGGGGGATG 301

Qy     462 agctagaaggcctgctggaagagaggcctgccctgtgcctgtccccgcaggccccatttc 521
           ||||    |||||| |||||||||||| ||||  |||| |||||||||||  ||  | |
Db     302 AGCT---CGGCCTGGTGGAAGAGAGGCCGGCCCATTGCCCGTCCCCGCAGGTGCCGGTGC 358

Qy     522 ccaagctgggctgggatgacgaactgcggaaacccggcgcccagatctacatgcgcttca 581
            || ||| |||||||| || || |||||||||||  | || ||| |||||||| |||||
Db     359 TCATGCTCGGCTGGGACGATGAGCTGCGGAAACCATGGGCTCAGGTCTACATGCACTTCA 418

Qy     582 tgcaggagcacacctgctacgatgccatggcaactagctccaagctagtcatcttcgaca 641
           |||||||||||||||||||||||||||||||||||| |||||||||| ||||||||||||
Db     419 TGCAGGAGCACACCTGCTACGATGCCATGGCAACCAGCTCCAAGCTGGTCATCTTCGACA 478

Qy     642 ccatgctggagatcaagaaggccttctttgctctggtggccaacggtgtgcgggcagccc 701
           |||||||  |||||||||||||||||||||| |||||||||| ||||| ||||| || |
Db     479 CCATGCTACAGATCAAGAAGGCCTTCTTTGCCCTGGTGGCCAATGGTGTCCGGGCCGCAC 538

Qy     702 ctctatggga 711
           ||||||||||
Db     539 CTCTATGGGA 548


RESULT   9
AL552459
LOCUS       AL552459                908 bp    mRNA    linear   EST 16-FEB-2001
DEFINITION  AL552459 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI070YC03 5
            prime, mRNA sequence.
ACCESSION   AL552459
VERSION     AL552459.1  GI:12891378
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 908)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .908
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DI070YC03"
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer.  Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      239 a    222 c    199 g    247 t      1 others
ORIGIN


  Query Match             21.3%;  Score 350;  DB 9;  Length 908;
  Best Local Similarity   65.6%;  Pred. No. 6.7e-72;
  Matches  509;  Conservative    1;  Mismatches  266;  Indels    0;  Gaps    0;

Qy     567 tctacatgcgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagc 626
           | || |    ||||||| ||   ||   |||||| ||    ||  | || ||||||||
Db     131 TGTATACTTCCTTCATGAAGTCTCATCGCTGCTATGACCTGATTCCCACAAGCTCCAAAT 190

Qy     627 tagtcatcttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacg 686
           | ||  | || || ||    ||| || | ||||| || || |||||| ||||| | ||||
Db     191 TGGTTGTATTTGATACGTCCCTGCAGGTGAAGAAAGCTTTTTTTGCTTTGGTGACTAACG 250

Qy     687 gtgtgcgggcagcccctctatgggacagcaagaagcagagctttgtggggatgctgacca 746
           |||| || || |||||| ||||||| || |||||||| || |||||||| ||||||||||
Db     251 GTGTACGAGCTGCCCCTTTATGGGATAGTAAGAAGCAAAGTTTTGTGGGCATGCTGACCA 310

Qy     747 tcactgacttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatg 806
           ||||||| ||||||    | ||||| |||||||| |  ||  || |||| ||||||||||
Db     311 TCACTGATTTCATCAATATCCTGCACCGCTACTATAAATCAGCCTTGGTACAGATCTATG 370

Qy     807 agattgaacaacataagattgagacctggagggagatctacctgcaaggctgcttcaagc 866
           || | ||| |||| ||||| || || ||||| ||| | || || || | || ||| || |
Db     371 AGCTAGAAGAACACAAGATAGAAACTTGGAGAGAGGTGTATCTCCAGGACTCCTTTAAAC 430

Qy     867 ctctggtctccatctctcctaatgataggctgtttgaagctgtctacaccctcatcaaga 926
           | || |||| ||| |||||||||  ||| ||||||| |||||||   |  | ||   ||
Db     431 CGCTTGTCTGCATTTCTCCTAATGCCAGCTTGTTTGATGCTGTCTCTTCATTAATTCGGA 490

Qy     927 accggatccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctca 986
           ||  ||||||  | ||||| ||| |||||||  |  ||||||||      |  |||||||||
Db     491 ACAAGATCCACAGGCTGCCAGTTATTGACCCAGAATCAGGCAATACTTTGTACATCCTCA 550

Qy     987 cacacaaacgcctgctcaagttcctgcacatctttggttccctgctgccccggccctcct 1046
           | ||||| ||| | || |||||||||  |  | |||     |   | | |||  |||    |
Db     551 CCCACAAGCGCATTCTGAAGTTCCTCAAATTGTTTATCACTGAGTTCCCCAAGCCAGAGT 610

Qy    1047 tcctctaccgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgc 1106
           || | | |    || |  |||| |    || ||||| |       |  | ||| |||| |
Db     611 TCATGTCCAAGTCTCTGGAAGAGCTACAGATTGGCACCTATGCCAATATTGCTATGGTTC 670

Qy    1107 tggagacagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcac 1166
              ||  | ||| ||      || |||| || |||||  | |  || || || || |
Db     671 GCACTACCACCCCCGTCTATGTGGCTCTGGGGATTTTTGTACAGCATCGAGTCTCAGCCC 730

Qy    1167 tgcctgtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattc 1226
           |||| |||||  | ||    || |  || |||| | |||| |||   ||||||||  ||
```

```
Db    731 TGCCAGTGGTGGATGAGAAGGGGCGTGTGGTGGACATCTACTCCAAGTTTGATGTTATCA 790

Qy   1227 acctggctgcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggc 1286
          | ||||| ||  |  | ||||||||| |||| ||  |   ||||    ||||| ||   |
Db    791 ATCTGGCAGCAGAAAAGACCTACAACAACCTAGATGTATCTGTGACTAAAGCCTTGCAAC 850

Qy   1287 agaggacactatgtctggagggagtcctttcctgccagccccacgagagcttgggg 1342
          |  |  |||  |   | ||||| || ||    ||| | |  || ||||   |||:|
Db    851 ATCGATCACATTACTTTGAGGGTGTTCTCAAGTGCTACCTGCATGAGACTCTGGRG 906


RESULT 10
BI858240
LOCUS       BI858240                 864 bp    mRNA    linear   EST 10-OCT-2001
DEFINITION  603384001F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5392777 5',
            mRNA sequence.
ACCESSION   BI858240
VERSION     BI858240.1  GI:15998987
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 864)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12001  row: a  column: 02
            High quality sequence stop: 830.
FEATURES             Location/Qualifiers
     source          1. .864
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5392777"
                     /clone_lib="NIH_MGC_87"
                     /tissue_type="mammary adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.383 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      223 a    202 c    203 g    236 t
ORIGIN


  Query Match             21.1%;  Score 346.8;  DB 10;  Length 864;
  Best Local Similarity   65.3%;  Pred. No. 3.7e-71;
  Matches  555;  Conservative    0;  Mismatches  292;  Indels    3;  Gaps    3;

Qy    608 atggcaactagctccaagctagtcatcttcgacaccatgctggagatcaagaaggccttc 667
          ||  | || ||||||||  | ||  | || || ||    ||| || | ||||| || ||
Db      2 ATTCCCACAAGCTCCAAATTGGTTGTATTTGATACGTCCCTGCAGGTGAAGAAAGCTTTT 61

Qy    668 tttgctctggtggccaacggtgtgcgggcagcccctctatgggacagcaagaagcagagc 727
          |||||| ||||| | |||||||| || || |||||| ||||||| || |||||||| ||
Db     62 TTTGCTTTGGTGACTAACGGTGTACGAGCTGCCCCTTTATGGGATAGTAAGAAGCAAAGT 121

Qy    728 tttgtggggatgctgaccatcactgacttcatcctggtgctgcatcgctactacaggtcc 787
          |||||||| |||||||||||||||| ||||||    | ||||| |||||||| |  ||
Db    122 TTTGTGGGCATGCTGACCATCACTGATTTCATCAATATCCTGCACCGCTACTATAAATCA 181

Qy    788 cccctggtccagatctatgagattgaacaacataagattgagacctggagggagatctac 847
           || |||| |||||||||||| | ||| |||| ||||| || || ||||| ||| | ||

Db    182 GCCTTGGTACAGATCTATGAGCTAGAAGAACACAAGATAGAAACTTGGAGAGAGGTGTAT 241

Qy    848 ctgcaaggctgcttcaagcctctggtctccatctctcctaatgatagcctgtttgaagct 907
          || || | || ||| || || || |||| ||| |||||||||  ||| |||||| |||
Db    242 CTCCAGGACTCCTTTAAACCGCTTGTCTGCATTTCTCCTAATGCCAGCTTGTTTGATGCT 301

Qy    908 gtctacaccctcatcaagaaccggatccatcgcctgcctgttcttgacccggtgtcaggc 967
          ||||   | | ||   |||| |||||| | ||||| ||| ||||||| |  ||||||
Db    302 GTCTCTTCATTAATTCGGAACAAGATCCACAGGCTGCCAGTTATTGACCCAGAATCAGGC 361

Qy    968 aacgtactccacatcctcacacacaaacgcctgctcaagttcctgcacatctttggttcc 1027
          ||     |  |||||||||| ||||| ||| | || ||||||||  |  | |||    |
Db    362 AATACTTTGTACATCCTCACCCACAAGCGCATTCTGAAGTTCCTCAAATTGTTTATCACT 421

Qy   1028 ctgctgccccggccctccttcctctaccgcactatccaagatttgggcatcggcacattc 1087
          | | |||  |||    ||| | | |    || |  |||  |    || |||||  |
Db    422 GAGTTCCCCAAGCCAGAGTTCATGTCCAAGTCTCTGGAAGAGCTACAGATTGGCACCTAT 481

Qy   1088 cgagacttggctgtggtgctggagacagcacccatcctgactgcactggacatctttgtg 1147
          |  | ||| |||| |     ||  | ||| ||      || |||| || |||||
Db    482 GCCAATATTGCTATGGTTCGCACTACCACCCCCGTCTATGTGGCTCTGGGGATTTTTGTA 541

Qy   1148 gaccggcgtgtgtctgcactgcctgtggtcaacgaatgtggtcaggtcgtgggcctctat 1207
          | |  || || || || ||||| |||||  | ||    || |  || |||| | ||||
Db    542 CAGCATCGAGTCTCAGCCCTGCCAGTGGTGGATGAGAAGGGGCGTGTGGTGGACATCTAC 601

Qy   1208 tcccgctttgatgtgattcacctggctgcccagcaaacctacaaccacctggacatgagt 1267
          |||   |||||||| ||  | |||| ||  |  | |||||||| |||| ||  |   |
Db    602 TCCAAGTTTGATGTTATCAATCTGGCAGCAGAAAAGACCTACAACAACCTAGATGTATCT 661

Qy   1268 gtgggagaagccctgaggcagaggacactatgtctggagggagtcctttcctgccagccc 1327
          |||   |||| ||   ||  |  ||| |   | ||||| || ||    ||| | |
Db    662 GTGACTAAAGCCTTGCAACATCGATCACATTACTTTGAGGGTGT-CTCAAGTGCTACCTG 720

Qy   1328 cacgagagcttgggggaagtgatcgacaggattg-ctcgggagcaggtacacaggctggt 1386
          || ||||   ||| |    | ||| ||||| | |     |  |  ||| ||| | || ||
Db    721 CATGAGACTCTGGAGACCATCATCAACAGGCTAGTGGAAGCCGAGGGTTCACCGACTTGT 780

Qy   1387 gctagtggacgagacccagcatctcttgggcgtggtctccctctccgacatccttcaggc 1446
          | ||||| || |   |     ||  |||    || || || || ||||||| |||||
Db    781 AGTGGTGGATGAAAATGATGTGGTCCAGGGACATGT-TCACTGTCTGACATCCTGCAGGC 839

Qy   1447 actggtgctc 1456
          |||||||||
Db    840 CCTGGTGCTC 849


RESULT 11
AL519198
LOCUS       AL519198                 826 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL519198 LTI_NFL011_NBC1 Homo sapiens cDNA clone CS0DA012YB06 5
            prime, mRNA sequence.
ACCESSION   AL519198
VERSION     AL519198.1  GI:12782691
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 826)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .826
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DA012YB06"
                     /clone_lib="LTI_NFL011_NBC1"
```

```
                     /sex="male"
                     /tissue_type="neuroblastoma cells"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library is not normalized, but is the control for
                     the normalized libraries. Library was constructed by Life
                     Technologies. Contact : Feng Liang Life Technologies, a
                     division of Invitrogen 9800 Medical Center Drive Rockville
                     , Maryland 20850, USA Fax : (1) 301 610 8371 Email :
                     fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT       217 a    202 c    175 g    232 t
ORIGIN


  Query Match             20.8%;  Score 343;  DB 9;  Length 826;
  Best Local Similarity   65.7%;  Pred. No. 2.9e-70;
  Matches  499;  Conservative    0;  Mismatches  260;  Indels    0;  Gaps    0;

Qy     567 tctacatgcgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagc 626
           | || |    ||||||| ||   ||   |||||| ||    ||  | || |||||||||
Db      68 TGTATACTTCCTTCATGAAGTCTCATCGCTGCTATGACCTGATTCCCACAAGCTCCAAAT 127

Qy     627 tagtcatcttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacg 686
           | ||  | || || ||    ||| || | ||||| || || |||||| ||||| | ||||
Db     128 TGGTTGTATTTGATACGTCCCTGCAGGTGAAGAAAGCTTTTTTTGCTTTGGTGACTAACG 187

Qy     687 gtgtgcgggcagcccctctatgggacagcaagaagcagagctttgtggggatgctgacca 746
           |||| || || |||||| |||||||| || |||||||| || |||||||| ||||||||||
Db     188 GTGTACGAGCTGCCCCTTTATGGGATAGTAAGAAGCAAAGTTTTGTGGGCATGCTGACCA 247

Qy     747 tcactgacttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatg 806
           ||||||| ||||||    | |||| |||||||| |  ||  || |||| ||||||||||
Db     248 TCACTGATTTCATCAATATCCTGCACCGCTACTATAAATCAGCCTTGGTACAGATCTATG 307

Qy     807 agattgaacaacataagattgagacctggagggagatctacctgcaaggctgcttcaagc 866
           || | ||| |||| ||||| || || ||||| ||| | || || || | || ||| || |
Db     308 AGCTAGAAGAACACAAGATAGAAACTTGGAGAGAGGTGTATCTCCAGGACTCCTTTAAAC 367

Qy     867 ctctggtctccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaaga 926
           | || |||| ||| |||||||||  ||| |||||| ||||||   |  | ||   ||
Db     368 CGCTTGTCTGCATTTCTCCTAATGCCAGCTTGTTTGATGCTGTCTCTTCATTAATTCGGA 427

Qy     927 accggatccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctca 986
           ||  ||||||  | ||||| ||| |||||||  |  ||||||||      |  |||||||||
Db     428 ACAAGATCCACAGGCTGCCAGTTATTGACCCAGAATCAGGCAATACTTTGTACATCCTCA 487

Qy     987 cacacaaacgcctgctcaagttcctgcacatctttggttccctgctgccccggccctcct 1046
           | ||||| ||| | || |||||||||  |  | |||     |   | | |||  |||    |
Db     488 CCCACAAGCGCATTCTGAAGTTCCTCAAATTGTTTATCACTGAGTTCCCCAAGCCAGAGT 547

Qy    1047 tcctctaccgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgc 1106
           || | | |    || |  ||||  |    || |||||  |       |  | ||| |||| |
Db     548 TCATGTCCAAGTCTCTGGAAGAGCTACAGATTGGCACCTATGCCAATATTGCTATGGTTC 607

Qy    1107 tggagacagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcac 1166
              ||  | ||| ||      || ||||  || |||||  |  |  || || || || |
Db     608 GCACTACCACCCCCGTCTATGTGGCTCTGGGGATTTTTGTACAGCATCGAGTCTCAGCCC 667

Qy    1167 tgcctgtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattc 1226
           |||| |||||  | ||    || |  || |||| | |||| |||   |||||||| ||
Db     668 TGCCAGTGGTGGATGAGAAGGGGCGTGTGGTGGACATCTACTCCAAGTTTGATGTTATCA 727

Qy    1227 acctggctgcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggc 1286
           | ||||| ||  |  | |||||||||| |||| ||  |  ||||   ||||| ||   |
Db     728 ATCTGGCAGCAGAAAAGACCTACAACAACCTAGATGTATCTGTGACTAAAGCCTTGCAAC 787

Qy    1287 agaggacactatgtctggagggagtccttttcctgccagc 1325
           |  |  |||  |   | ||||| || ||    ||| | |
Db     788 ATCGATCACATTACTTTGAGGGTGTTCTCAAGTGCTACC 826
```

```
RESULT  12
AL513784
LOCUS       AL513784                951 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL513784 LTI_NFL006_PL2 Homo sapiens cDNA clone CL0BA003ZH07 5
            prime, mRNA sequence.
ACCESSION   AL513784
VERSION     AL513784.1  GI:12777278
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 951)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .951
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CL0BA003ZH07"
                     /clone_lib="LTI_NFL006_PL2"
                     /tissue_type="placenta"
                     /note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed by
                     Life Technologies. Contact : Feng Liang Life Technologies,
                     a division of Invitrogen 9800 Medical Center Drive
                     Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
                     Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      295 a    190 c    213 g    252 t      1 others
ORIGIN


  Query Match             20.8%;  Score 342.8;  DB 9;  Length 951;
  Best Local Similarity   62.7%;  Pred. No. 3.4e-70;
  Matches  533;  Conservative    0;  Mismatches  317;  Indels    0;  Gaps    0;

Qy     567 tctacatgcgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagc 626
           | |||||||| ||||||  |   ||||  || || ||   | |  |||| || || ||||
Db     101 TTTACATGCGATTCATGAGGTCACACAAGTGTTATGACATCGTTCCAACCAGTTCAAAGC 160

Qy     627 tagtcatcttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacg 686
           | ||  |||| || || |   |  |  | || ||||||||||||||| |||| |||||||
Db     161 TTGTTGTCTTTGATACTACATTACAAGTTAAAAAGGCCTTCTTTGCTTTGGTAGCCAACG 220

Qy     687 gtgtgcgggcagcccctctatgggacagcaagaagcagagctttgtggggatgctgacca 746
           |||| || |||||  || || |||||  || || || || || |||||  || |||||  || |
Db     221 GTGTCCGAGCAGCGCCACTGTGGGAGAGTAAAAAACAAAGTTTTGTAGGAATGCTAACAA 280

Qy     747 tcactgacttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatg 806
           | || || |||||      | || ||| | |||||  |  || ||  |||| |||||  ||||
Db     281 TTACAGATTTCATAAATATACTACATAGATACTATAAATCACCTATGGTACAGATTTATG 340

Qy     807 agattgaacaacataagattgagacctggagggagatctacctgcaaggctgcttcaagc 866
           |  | ||  ||||||| ||||| || |||||||||  | ||  | ||||     || ||||
Db     341 AATTAGAGGAACATAAAATTGAAACATGGAGGGAGCTTTATTTACAAGAAACATTTAAGC 400

Qy     867 ctctggtctccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaaga 926
           || | ||    || |||||  |||  ||||| || || ||||| ||| || | |||||  |
Db     401 CTTTAGTGAATATATCTCCAGATGCAAGCCTCTTCGATGCTGTATACTCCTTGATCAAAA 460

Qy     927 accggatccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctca 986
           |    |||||  |  |||| ||| |||||||  |    || || | |||  | || || |
```

```
Db     461 ATAAAATCCACAGATTGCCCGTTATTGACCCTATCAGTGGGAATGCACTTTATATACTTA 520

Qy     987 cacacaaacgcctgctcaagttcctgcacatctttggttccctgctgccccggccctcct 1046
           | |||||| |  | ||||||||||| ||  | |||   ||     ||||  |||  |||
Db     521 CCCACAAAAGAATCCTCAAGTTCCTCCAGCTTTTTATGTCTGATATGCCAAAGCCTGCCT 580

Qy    1047 tcctctaccgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgc 1106
           || |  | |  |   |  | ||  | || || || || | ||   || | ||  |  | |
Db     581 TCATGAAGCAGAACCTGGATGAGCTTGGAATAGGAACGTACCACAACATTGCCTTCATAC 640

Qy    1107 tggagacagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcac 1166
                 | |||||| | |  ||  || |||| ||||||||  | ||  | || || |
Db     641 ATCCAGACACTCCCATCATCAAAGCCTTGAACATATTTGTGGAAAGACGAATATCAGCTC 700

Qy    1167 tgcctgtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattc 1226
           ||||||| ||  | || |  ||  | || || |   | ||||||  ||||||| |||
Db     701 TGCCTGTTGTGGATGAGTCAGGAAAAGTTGTAGATATTTATTCCAAATTTGATGTAATTA 760

Qy    1227 acctggctgcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggc 1286
           | || |||||  || |||| |||||  |||| || || |  |||    | |||||  ||
Db     761 ATCTTGCTGCTGAGAAAACATACAATAACCTAGATATCACGGTGACCCAGGCCCTTCAGC 820

Qy    1287 agaggacactatgtctggagggagtcctttcctgccagccccacgagagcttgggggaag 1346
           |  |  ||| | | | || || ||  |    ||| |    |  || |   ||| |
Db     821 ACCGTTCACAGTATTTTGAAGGTGTTGTGAAGTGCAATAAGCTGGAAATACTGGAGACCA 880

Qy    1347 tgatcgacaggattgctcgggagcaggtacacaggctggtgctagtggacgagacccagc 1406
           |  | ||||| || |   | |   |||| ||  |||||||| | ||  | ||  |   |
Db     881 TCGTCGACAGAATAGTAAGAGCTGAGGTCCATCGGCTGGTGGTGGTAAATGAAGCAGAWA 940

Qy    1407 atctcttggg 1416
             | |  ||||
Db     941 GTATTGTGGG 950


RESULT 13
BM462694
LOCUS       BM462694                1013 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6427350 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5540312
            5', mRNA sequence.
ACCESSION   BM462694
VERSION     BM462694.1  GI:18511734
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1013)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12236  row: d  column: 09
            High quality sequence stop: 640.
FEATURES             Location/Qualifiers
     source          1. .1013
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5540312"
                     /clone_lib="NIH_MGC_72"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                                Technologies."
BASE COUNT      272 a    241 c    224 g    275 t      1 others
ORIGIN


  Query Match             20.4%;  Score 336.6;  DB 10;  Length 1013;
  Best Local Similarity   65.5%;  Pred. No. 1e-68;
  Matches  492;  Conservative    0;  Mismatches  259;  Indels    0;  Gaps    0;

Qy     567 tctacatgcgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagc 626
           | || |    ||||||| ||   ||   |||||| ||    ||  | || ||||||||
Db      70 TGTATACTTCCTTCATGAAGTCTCATCGCTGCTATGACCTGATTCCCACAAGCTCCAAAT 129

Qy     627 tagtcatcttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacg 686
           | ||  | || || ||    ||| || | ||||| || || |||||| ||||| | ||||
Db     130 TGGTTGTATTTGATACGTCCCTGCAGGTGAAGAAAGCTTTTTTTGCTTTGGTGACTAACG 189

Qy     687 gtgtgcgggcagcccctctatgggacagcaagaagcagagctttgtggggatgctgacca 746
           |||| || || |||||| ||||||| || |||||||| || |||||||| ||||||||||
Db     190 GTGTACGAGCTGCCCCTTTATGGGATAGTAAGAAGCAAAGTTTTGTGGGCATGCTGACCA 249

Qy     747 tcactgacttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatg 806
           ||||||| ||||||    | ||||| |||||||| |  ||  || |||| ||||||||||
Db     250 TCACTGATTTCATCAATATCCTGCACCGCTACTATAAATCAGCCTTGGTACAGATCTATG 309

Qy     807 agattgaacaacatagattgagacctggagggagatctacctgcaaggctgcttcaagc 866
           || | ||| |||| ||||| || || ||||| ||| | || || || | || ||| || |
Db     310 AGCTAGAAGAACACAAGATAGAAACTTGGAGAGAGGTGTATCTCCAGGACTCCTTTAAAC 369

Qy     867 ctctggtctccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaaga 926
           | || |||| ||| |||||||||  ||| |||||||| |||||||   |  | ||  ||
Db     370 CGCTTGTCTGCATTTCTCCTAATGCCAGCTTGTTTGATGCTGTCTCTTCATTAATTCGGA 429

Qy     927 accggatccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctca 986
           ||  ||||||  | ||||| ||| ||||||| |  |||||||      |  |||||||||
Db     430 ACAAGATCCACAGGCTGCCAGTTATTGACCCAGAATCAGGCAATACTTTGTACATCCTCA 489

Qy     987 cacacaaacgcctgctcaagttcctgcacatctttggttccctgctgccccggccctcct 1046
           | ||||| ||| | || ||||||||  |  | |||    |   | | |||  |||    |
Db     490 CCCACAAGCGCATTCTGAAGTTCCTCAAATTGTTTATCACTGAGTTCCCCAAGCCAGAGT 549

Qy    1047 tcctctaccgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgc 1106
           || | | |    || |  |||| |    || ||||| |       |  | ||| |||| |
Db     550 TCATGTCCAAGTCTCTGGAAGAGCTACAGATTGGCACCTATGCCAATATTGCTATGGTTC 609

Qy    1107 tggagacagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcac 1166
              ||  | ||| ||       || ||||  || |||||  | |  || || || || |
Db     610 GCACTACCACCCCCGTCTATGTGGCTCTGGGGATTTTTGTACAGCATCGAGTCTCAGCCC 669

Qy    1167 tgcctgtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattc 1226
           |||| |||||  | ||    || |  || |||| | |||| |||   ||||||| ||
Db     670 TGCCAGTGGTGGATGAGAAAGGGCGTGTGGTGGACATCTACTCCAAGTTTGATGTTATCA 729

Qy    1227 acctggctgcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggc 1286
           | ||||| ||  |  | |||||||||| |||| ||  |  ||||    ||||| ||  |
Db     730 ATCTGGCAGCAGAAAAGACCTACAACAACCTAGATGTATCTGTGACTAAAGCCTTGCAAC 789

Qy    1287 agaggacactatgtctggagggagtcctttc 1317
           |  |  |||  |   |    || ||  | ||
Db     790 ATCGATCACATTACTTTTGAGGGGTGGTCTC 820


RESULT 14
BM470188
LOCUS       BM470188                1004 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6476961 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5553194
            5', mRNA sequence.
ACCESSION   BM470188
VERSION     BM470188.1  GI:18519230
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1  (bases 1 to 1004)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Lou Staudt
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12269  row: m  column: 03
            High quality sequence stop: 745.
FEATURES             Location/Qualifiers
     source          1. .1004
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5553194"
                     /clone_lib="NIH_MGC_85"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.867 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      254 a    247 c    232 g    265 t      6 others
ORIGIN
```

Query Match 20.3%; Score 335; DB 10; Length 1004;
Best Local Similarity 66.1%; Pred. No. 2.4e-68;
Matches 482; Conservative 0; Mismatches 247; Indels 0; Gaps 0;

```
Qy     567 tctacatgcgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagc 626
           | || |    ||||||| ||   ||   |||||| ||    ||  | || ||||||||
Db     133 TGTATACTTCCTTCATGAAGTCTCATCGCTGCTATGACCTGATTCCCACAAGCTCCAAAT 192

Qy     627 tagtcatcttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacg 686
           | ||  | || || ||    ||| || | ||||| || || |||||| ||||| | ||||
Db     193 TGGTTGTATTTGATACGTCCCTGCAGGTGAAGAAAGCTTTTTTTGCTTTGGTGACTAACG 252

Qy     687 gtgtgcgggcagcccctctatgggacagcaagaagcagagctttgtggggatgctgacca 746
           |||| || || |||||| ||||||| || |||||||| || |||||||| ||||||||||
Db     253 GTGTACGAGCTGCCCCTTTATGGGATAGTAAGAAGCAAAGTTTTGTGGGCATGCTGACCA 312

Qy     747 tcactgacttcatcctggtgctgcatcgctactacaggtccccctggtccagatctatg 806
           ||||||| ||||||    | ||||| |||||||| |  ||  || |||| ||||||||||
Db     313 TCACTGATTTCATCAATATCCTGCACCGCTACTATAAATCAGCCTTGGTACAGATCTATG 372

Qy     807 agattgaacaacataagattgagacctggagggagatctacctgcaaggctgcttcaagc 866
           || | ||| |||| ||||| || || ||||| ||| | || || || | || ||| || |
Db     373 AGCTAGAAGAACACAAGATAGAAACTTGGAGAGAGGTGTATCTCCAGGACTCCTTTAAAC 432

Qy     867 ctctggtctccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaaga 926
           | || |||| ||| |||||||||  ||| ||||||| |||||||   |  | ||   ||
Db     433 CGCTTGTCTGCATTTCTCCTAATGCCAGCTTGTTTGATGCTGTCTCTTCATTAATTCGGA 492

Qy     927 accggatccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctca 986
           ||  ||||||  | ||||| ||| |||||||  |  ||||||||     |  |||||||||
Db     493 ACAAGATCCACAGGCTGCCAGTTATTGACCCAGAATCAGGCAATACTTTGTACATCCTCA 552

Qy     987 cacacaaacgcctgctcaagttcctgcacatctttggttccctgctgccccggccctcct 1046
           | ||||| ||| | || |||||||| |  | |||    |   | | |||  |||    |
Db     553 CCCACAAGCGCATTCTGAAGTTCCTCAAATTGTTTATCACTGAGTTCCCCAAGCCAGAGT 612

Qy    1047 tcctctaccgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgc 1106
           || | | |    || |  |||| |    || ||||| |       |  | ||| |||| |
Db     613 TCATGTCCAAGTCTCTGGAAGAGCTACAGATTGGCACCTATGCCAATATTGCTATGGTTC 672
```

```
Qy    1107 tggagacagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcac 1166
           ||  | ||| ||      || |||| || |||||  | |  || || || || | |
Db     673 GCACTACCACCCCCGTCTATGTGGCTCTGGGGATTTTTGTACAGCATCGAGTCTCAGCCC 732

Qy    1167 tgcctgtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattc 1226
           |||| |||||  | ||    || |  || |||| | |||| || |  |||||||| ||
Db     733 TGCCAGTGGTGGATGAGAAGGGGCGTGTGGTGGACATCTACTCNCAGTTTGATGTTATCA 792

Qy    1227 acctggctgcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggc 1286
           | ||||| ||  |  | |||||||||| |||| ||  |   ||||     |||| ||   |
Db     793 ATCTGGCAGCAGAAAAGACCTACAACAACCTAGATGTATCTGTGACTANAGCCTTGCAAC 852

Qy    1287 agaggacac 1295
           |  |  |||
Db     853 ATCGATCAC 861
```

```
RESULT 15
BM487789
LOCUS       BM487789                595 bp    mRNA    linear   EST 07-FEB-2002
DEFINITION  pgm2n.pk005.j24 Normalized Chicken Breast Muscle, Leg Muscle, and
            Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cDNA
            clone pgm2n.pk005.j24 5' similar to gb|AAC52580.1 (U42413)
            5'-AMP-activated protein kinase, gamma-1 subunit [Rattus norvegicus
            ], mRNA sequence.
ACCESSION   BM487789
VERSION     BM487789.1  GI:18608720
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 595)
  AUTHORS   Cogburn,L.A. and Monsonego-Ornan,E.
  TITLE     ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and
            Epiphyseal Growth Plate cDNA library, USDA/IFAFS Animal Genome
            Project
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Larry A. Cogburn
            University of Delaware
            Townsend Hall, Newark, DE 19717, USA
            Tel: 302-831-1335
            Fax: 302-831-2822
            Email: cogburn@udel.edu, www.chickest.udel.edu.
FEATURES             Location/Qualifiers
     source          1. .595
                     /organism="Gallus gallus"
                     /strain="Commercial broiler and  Ottawa Res. Centre
                     Strains 90 & 21"
                     /db_xref="taxon:9031"
                     /clone="pgm2n.pk005.j24"
                     /clone_lib="Normalized Chicken Breast Muscle, Leg Muscle,
                     and Epiphyseal Growth Plate cDNA library (pgm2n)"
                     /sex="Male and Female"
                     /tissue_type="Breast muscle, leg muscle and epiphyseal
                     growth plate"
                     /dev_stage="Breast,leg:Embryo(d19);post-hatch(1d,1,3,5,7,9
                     ,11 weeks);growth plate(1d,7d,14d post-hatch)"
                     /lab_host="E. coli EMDH10B"
                     /note="Vector: pCMVSPORT6; Library made from equivalent
                     pools of total RNA isolated from each tissue (embryonic
                     muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth
                     plate 33.3% of the final RNA pool). Single pass sequencing
                     from 5'-end"
BASE COUNT      126 a    199 c    137 g    133 t
ORIGIN
```

Query Match 20.0%; Score 329.2; DB 10; Length 595;
Best Local Similarity 72.4%; Pred. No. 4.4e-67;
Matches 427; Conservative 0; Mismatches 163; Indels 0; Gaps 0;

```
Qy     595 ctgctacgatgccatggcaactagctccaagctagtcatcttcgacaccatgctggagat 654
           |||||||||   |||  | || |||||||| || ||| ||||||||||    ||  || |
Db       4 CTGCTACGACCTCATCCCCACCAGCTCCAAACTCGTCGTCTTCGACACTTCCCTTCAGGT 63

Qy     655 caagaaggcctttctttgctctggtggccaacggtgtgcgggcagcccctctatgggacag 714
            |||||||| |||||||| |||||  |||| || |||||||||||||| || ||||||||
Db      64 GAAGAAGGCTTTCTTTGCACTGGTCACCAATGGCGTGCGGGCAGCCCCGCTGTGGGACAG 123

Qy     715 caagaagcagagctttgtggggatgctgaccatcactgacttcatcctggtgctgcatcg 774
           ||| ||||||||||||||||| ||||||||||||||||||||||||    | ||||| ||
Db     124 CAAAAAGCAGAGCTTTGTGGGCATGCTGACCATCACTGACTTCATCAACATCCTGCACCG 183

Qy     775 ctactacaggtcccccctggtccagatctatgagattgaacaacataagattgagacctg 834
           |||||||| ||| ||| |||| |||||||||||| | ||  | || || || ||||| ||
Db     184 CTACTACAAGTCGCCCATGGTGCAGATCTATGAGCTGGAGGAGCACAAAATCGAGACGTG 243

Qy     835 gagggagatctacctgcaaggctgcttcaagcctctggtctccatctctcctaatgatag 894
           ||| ||| | |||||||| | || |||||||||  |||||| |||||| || ||||  ||
Db     244 GAGAGAGGTGTACCTGCAGGACTCCTTCAAGCCGTTGGTCTGCATCTCCCCCAATGCCAG 303

Qy     895 cctgtttgaagctgtctacaccctcatcaagaaccggatccatcgcctgcctgttcttga 954
           ||| ||||| || |||| | |||| |||   ||   |||||| ||||| || ||  | ||
Db     304 CCTTTTTGATGCCGTCTCCTCCCTGATCCGCAATAAGATCCACCGCCTCCCCGTCATCGA 363

Qy     955 cccggtgtcaggcaacgtactccacatcctcacacacaaacgcctgctcaagttcctgca 1014
           ||| |  || ||||||   ||| |||||||||| |||||||| | ||||||||||  |
Db     364 CCCCGACTCGGGCAACACTCTCTACATCCTCACCCACAAACGCATCCTCAAGTTCCTCAA 423

Qy    1015 catctttggttccctgctgccccggccctccttcctctaccgcactatccaagatttggg 1074
             |||||  | |   | | ||  |||   ||  |  ||| ||| |  | ||  ||
Db     424 ACTCTTTATTGCAGAGGTCCCAAAGCCTGAGTTTATGGCCCGGACTTTGGAGGAGCTGCA 483

Qy    1075 catcggcacattccgagacttggctgtggtgctggagacagcacccatcctgactgcact 1134
           |||||  || | | |  || | || ||||||      |   | ||||||     || ||
Db     484 GATCGGTACCTACAGCAACATCGCCGTGGTGAGCACCAGCACTCCCATCTATGTGGCTCT 543

Qy    1135 ggacatctttgtggaccggcgtgtgtctgcactgcctgtggtcaacgaat 1184
           || |||||| ||| | |  || || || || ||||| ||||||  | || |
Db     544 GGGCATCTTCGTGCAGCACCGCGTCTCCGCTCTGCCCGTGGTCGATGATT 593
```

Search completed: October 3, 2002, 16:15:01
Job time: 16956 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 16:21:45 ; Search time 180.77 Seconds
(without alignments)
2237.973 Million cell updates/sec

Title: US-09-826-581-5
Perfect score: 1647
Sequence: 1 ttggtctggggctggccaca.....acaccagcctcttagtcttc 1647

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

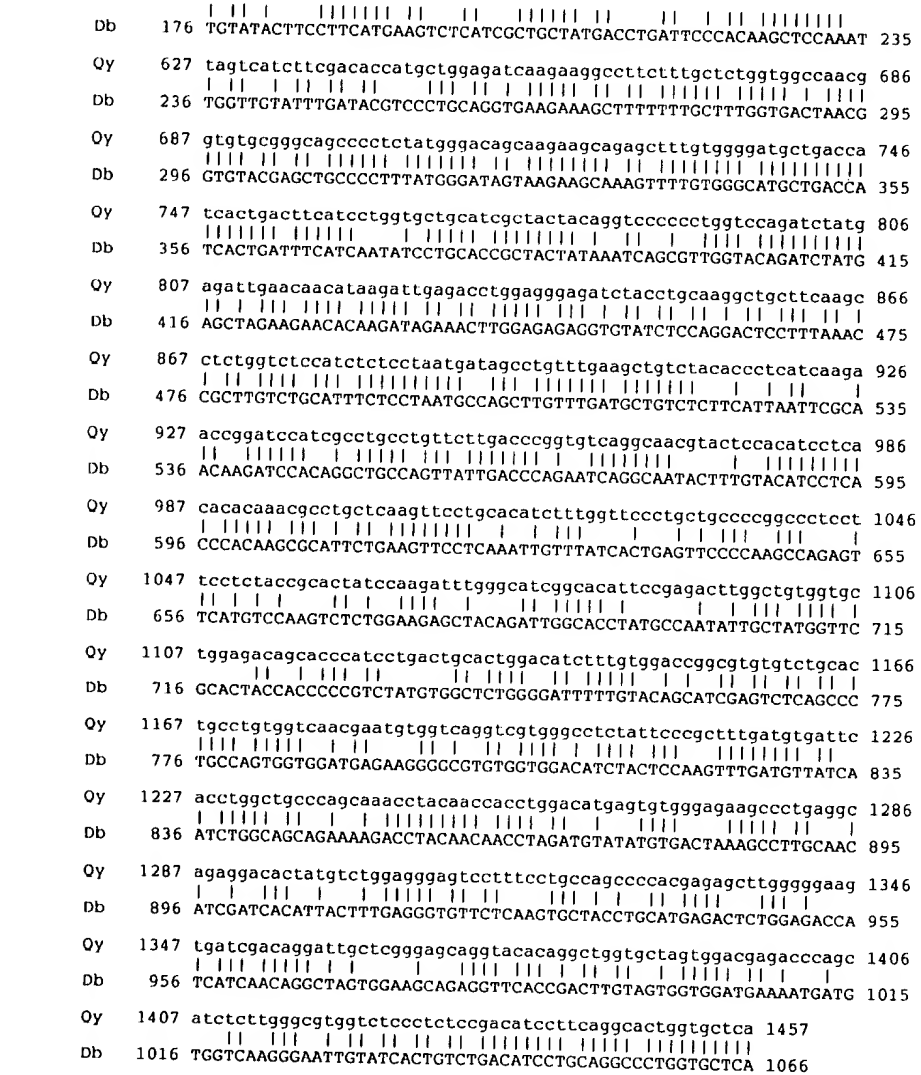
	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	390.2	23.7	1576	3	US-09-101-146-63	Sequence 63, Appl
	2	366.6	22.3	1435	2	US-08-878-989-14	Sequence 14, Appl
	3	366.6	22.3	1435	4	US-09-272-796-14	Sequence 14, Appl
	4	49	3.0	7218	1	US-08-232-463-14	Sequence 14, Appl
	5	41.2	2.5	1022	4	US-09-056-105-8	Sequence 8, Appli
c	6	40.2	2.4	7218	1	US-08-232-463-14	Sequence 14, Appl
	7	39.6	2.4	1022	4	US-09-056-105-18	Sequence 18, Appl
	8	39.6	2.4	2531	1	US-08-299-849B-14	Sequence 14, Appl
	9	39.6	2.4	2531	2	US-08-142-368A-14	Sequence 14, Appl
	10	39.6	2.4	2531	3	US-08-967-727-14	Sequence 14, Appl
	11	39.6	2.4	2531	4	US-08-037-230D-14	Sequence 14, Appl
	12	39.6	2.4	4895	4	US-09-056-105-10	Sequence 10, Appl
c	13	38.8	2.4	289	4	US-09-007-005-17	Sequence 17, Appl
c	14	38.8	2.4	289	4	US-09-244-796-17	Sequence 17, Appl
	15	38.8	2.4	6816	4	US-09-404-650-1	Sequence 1, Appli
	16	38.8	2.4	6855	4	US-09-404-650-3	Sequence 3, Appli
	17	38.4	2.3	1458	4	US-09-500-569-15	Sequence 15, Appl
c	18	38.2	2.3	3489	2	US-08-728-323A-1	Sequence 1, Appli
	19	38.2	2.3	32207	2	US-08-770-379-20	Sequence 20, Appl
	20	38.2	2.3	32207	4	US-08-757-669A-20	Sequence 20, Appl
	21	38.2	2.3	32207	4	US-09-230-371A-20	Sequence 20, Appl
	22	38	2.3	1931	2	US-09-130-114-2	Sequence 2, Appli
	23	38	2.3	2531	1	US-08-299-849B-13	Sequence 13, Appl
	24	38	2.3	2531	2	US-08-142-368A-13	Sequence 13, Appl
	25	38	2.3	2531	3	US-08-967-727-13	Sequence 13, Appl
	26	38	2.3	2531	4	US-08-037-230D-13	Sequence 13, Appl
	27	38	2.3	11495	4	US-09-056-105-9	Sequence 9, Appli
c	28	36.8	2.2	2580	3	US-09-050-863-2	Sequence 2, Appli
c	29	36.8	2.2	2580	4	US-09-359-081-2	Sequence 2, Appli
	30	36.8	2.2	5452	2	US-09-130-114-1	Sequence 1, Appli
c	31	36.8	2.2	9600	4	US-08-910-647-1	Sequence 1, Appli
c	32	36.8	2.2	10596	1	US-07-884-811-15	Sequence 15, Appl
c	33	36.8	2.2	10596	1	US-07-885-971-15	Sequence 15, Appl
c	34	36.8	2.2	10596	1	US-08-087-783A-15	Sequence 15, Appl
c	35	36.8	2.2	10596	1	US-08-194-088B-15	Sequence 15, Appl
c	36	36.8	2.2	10596	2	US-08-194-087-15	Sequence 15, Appl
c	37	36.8	2.2	10596	5	PCT-US93-04648-15	Sequence 15, Appl
	38	36.8	2.2	12492	6	5206163-2	Patent No. 5206163
	39	36.4	2.2	51259	3	US-08-781-891-209	Sequence 209, App
c	40	36.2	2.2	1900	4	US-09-608-285A-47	Sequence 47, Appl
c	41	36.2	2.2	2294	4	US-09-608-285A-49	Sequence 49, Appl
c	42	36.2	2.2	2371	4	US-09-608-285A-46	Sequence 46, Appl
c	43	36.2	2.2	2497	4	US-09-608-285A-51	Sequence 51, Appl
c	44	36.2	2.2	2693	4	US-09-608-285A-48	Sequence 48, Appl
c	45	36.2	2.2	2762	4	US-09-608-285A-26	Sequence 26, Appl

ALIGNMENTS

```
RESULT 1
US-09-101-146-63
; Sequence 63, Application US/09101146
; Patent No. 6124125
;  GENERAL INFORMATION:
;    APPLICANT:  Dartmouth College, St. Vincents Institute of
;    APPLICANT:  Medical Research, Kemp et al.
;    TITLE OF INVENTION:  No. 6124125el AMP Activated Protein Kinase
;    NUMBER OF SEQUENCES:  64
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Jane Massey Licata, Esq.
;      STREET:  66 E. Main Street
;      CITY:  Marlton
;      STATE:  NJ
;      COUNTRY:  USA
;      ZIP:  08053
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  DISKETTE, 3.5 INCH, 1.44 Mb STORAGE
;      COMPUTER:  IBM PC
;      OPERATING SYSTEM:  WINDOWS 95
;      SOFTWARE:  WORDPERFECT 6.0 FOR WINDOWS
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/101,146
;      FILING DATE:  October 7, 1998
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PN7450
;      FILING DATE:  8 JAN 1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Jane Massey Licata
;      REGISTRATION NUMBER:  32,257
;      REFERENCE/DOCKET NUMBER:  DC-0050
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (856) 810-1515
;      TELEFAX:  (856) 810-1454
;  INFORMATION FOR SEQ ID NO:  63:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1576
;      TYPE:  Nucleic acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
;    ANTI-SENSE:  No
US-09-101-146-63


  Query Match           23.7%;  Score 390.2;  DB 3;  Length 1576;
  Best Local Similarity  64.9%;  Pred. No. 1.8e-93;
  Matches  578;  Conservative    0;  Mismatches  313;  Indels    0;  Gaps    0;

Qy    567 tctacatgcgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagc 626
```



```
Db    176 TGTATACTTCCTTCATGAAGTCTCATCGCTGCTATGACCTGATTCCCACAAGCTCCAAAT 235

Qy    627 tagtcatcttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacg 686

Db    236 TGGTTGTATTTGATACGTCCCTGCAGGTGAAGAAAGCTTTTTTTGCTTTGGTGACTAACG 295

Qy    687 gtgtgcgggcagcccctctatgggacagcaagaagcagagctttgtggggatgctgacca 746

Db    296 GTGTACGAGCTGCCCCTTTATGGGATAGTAAGAAGCAAAGTTTTGTGGGCATGCTGACCA 355

Qy    747 tcactgacttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatg 806

Db    356 TCACTGATTTCATCAATATCCTGCACCGCTACTATAAATCAGCGTTGGTACAGATCTATG 415

Qy    807 agattgaacaacataagattgagacctggagggagatctacctgcaaggctgcttcaagc 866

Db    416 AGCTAGAAGAACACAAGATAGAAACTTGGAGAGAGGTGTATCTCCAGGACTCCTTTAAAC 475

Qy    867 ctctggtctccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaaga 926

Db    476 CGCTTGTCTGCATTTCTCCTAATGCCAGCTTGTTTGATGCTGTCTCTTCATTAATTCGCA 535

Qy    927 accggatccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctca 986

Db    536 ACAAGATCCACAGGCTGCCAGTTATTGACCCAGAATCAGGCAATACTTTGTACATCCTCA 595

Qy    987 cacacaaacgcctgctcaagttcctgcacatctttggttccctgctgccccggccctcct 1046

Db    596 CCCACAAGCGCATTCTGAAGTTCCTCAAATTGTTTATCACTGAGTTCCCCAAGCCAGAGT 655

Qy   1047 tcctctaccgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgc 1106

Db    656 TCATGTCCAAGTCTCTGGAAGAGCTACAGATTGGCACCTATGCCAATATTGCTATGGTTC 715

Qy   1107 tggagacagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcac 1166

Db    716 GCACTACCACCCCCGTCTATGTGGCTCTGGGGATTTTTGTACAGCATCGAGTCTCAGCCC 775

Qy   1167 tgcctgtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattc 1226

Db    776 TGCCAGTGGTGGATGAGAAGGGGCGTGTGGTGGACATCTACTCCAAGTTTGATGTTATCA 835

Qy   1227 acctggctgcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggc 1286

Db    836 ATCTGGCAGCAGAAAAGACCTACAACAACCTAGATGTATATGTGACTAAAGCCTTGCAAC 895

Qy   1287 agaggacactatgtctggagggagtcctttcctgccagccccacgagagcttgggggaag 1346

Db    896 ATCGATCACATTACTTTGAGGGTGTTCTCAAGTGCTACCTGCATGAGACTCTGGAGACCA 955

Qy   1347 tgatcgacaggattgctcgggagcaggtacacaggctggtgctagtggacgagacccagc 1406

Db    956 TCATCAACAGGCTAGTGGAAGCAGAGGTTCACCGACTTGTAGTGGTGGATGAAAATGATG 1015

Qy   1407 atctcttgggcgtggtctccctctccgacatccttcaggcactggtgctca 1457

Db   1016 TGGTCAAGGGAATTGTATCACTGTCTGACATCCTGCAGGCCCTGGTGCTCA 1066
```

```
RESULT   2
US-08-878-989-14
; Sequence 14, Application US/08878989
; Patent No. 5885803
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Guegler, Karl G.
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  DISEASE ASSOCIATED PROTEIN
;    TITLE OF INVENTION:  KINASES
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/878,989
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J J
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0321 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1435 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  PENITUT01
;      CLONE:  1452972
US-08-878-989-14


  Query Match             22.3%;  Score 366.6;  DB 2;  Length 1435;
  Best Local Similarity   62.9%;  Pred. No. 2.8e-87;
  Matches  567;  Conservative    0;  Mismatches  334;  Indels    0;  Gaps    0;

Qy    567 tctacatgcgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagc 626

Db    183 TTTACATGCGATTCATGAGGTCACACAAGTGTTATGACATCGTTCCAACCAGTTCAAAGC 242

Qy    627 tagtcatcttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacg 686

Db    243 TTGTTGTCTTTGATACTACATTACAAGTTAAAAAGGCCTTCTTTGCTTTGGTAGCCAACG 302

Qy    687 gtgtgcgggcagcccctctatgggacagcaagaagcagagctttgtggggatgctgacca 746

Db    303 GTGTCCGAGCAGCGCCACTGTGGGAGAGTAAAAAACAAAGTTTTGTAGGAATGCTAACAA 362

Qy    747 tcactgacttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatg 806

Db    363 TTACAGATTTCATAAATATACTACATAGATACTATAAATCACCTATGGTACAGATTTATG 422

Qy    807 agattgaacaacataagattgagacctggagggagatctacctgcaaggctgcttcaagc 866

Db    423 AATTAGAGGAACATAAAATTGAAACATGGAGGGAGCTTTATTTACAAGAAACATTTAAGC 482

Qy    867 ctctggtctccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaaga 926

Db    483 CTTTAGTGAATATATCTCCAGATGCAAGCCTCTTCGATGCTGTATACTCCTTGATCAAAA 542

Qy    927 accggatccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctca 986

Db    543 ATAAAATCCACAGATTGCCCGTTATTGACCCTATCAGTGGGAATGCACTTTATATACTTA 602
```

```
Qy    987 cacacaaacgcctgctcaagttcctgcacatctttggttccctgctgccccggccctcct 1046
          | |||||| |  | ||||||||||| ||  | |||   ||     ||||  |||  |||
Db    603 CCCACAAAAGAATCCTCAAGTTCCTCCAGCTTTTTATGTCTGATATGCCAAAGCCTGCCT 662

Qy   1047 tcctctaccgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgc 1106
          || |  | |   |   | ||  | || || || || | ||   || | ||  |  | |
Db    663 TCATGAAGCAGAACCTGGATGAGCTTGGAATAGGAACGTACCACAACATTGCCTTCATAC 722

Qy   1107 tggagacagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcac 1166
                | |||||| | |  ||  || |||| ||||||||  | ||  | || || |
Db    723 ATCCAGACACTCCCATCATCAAAGCCTTGAACATATTTGTGGAAAGACGAATATCAGCTC 782

Qy   1167 tgcctgtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattc 1226
          |||||||  ||  | ||  |  ||  | || ||  |   | ||||||   |||||||| |||
Db    783 TGCCTGTTGTGGATGAGTCAGGAAAAGTTGTAGATATTTATTCCAAATTTGATGTAATTA 842

Qy   1227 acctggctgcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggc 1286
          | || |||||  || |||| |||||  |||| || || |  |||    | |||||   ||
Db    843 ATCTTGCTGCTGAGAAAACATACAATAACCTAGATATCACGGTGACCCAGGCCCTTCAGC 902

Qy   1287 agaggacactatgtctggagggagtcctttcctgccagccccacgagagcttgggggaag 1346
          |  |  |||  | | | || || ||  |    ||| |    |  || |   ||| |
Db    903 ACCGTTCACAGTATTTTGAAGGTGTTGTGAAGTGCAATAAGCTGGAAATACTGGAGACCA 962

Qy   1347 tgatcgacaggattgctcgggagcaggtacacaggctggtgctagtggacgagacccagc 1406
          |  | ||||| || |   | |   |||| ||  |||||||| | ||  | ||  |  |
Db    963 TCGTGGACAGAATAGTAAGAGCTGAGGTCCATCGGCTGGTGGTGGTAAATGAAGCAGATA 1022

Qy   1407 atctcttgggcgtggtctccctctccgacatccttcaggcactggtgctcagccctgctg 1466
           | |  |||| |  | ||||| || ||||| || || || ||| | ||||  || || |
Db   1023 GTATTGTGGGTATTATTTCCCTGTCGGACATTCTGCAAGCCCTGATCCTCACACCAGCAG 1082

Qy   1467 g 1467
          |
Db   1083 G 1083
```

```
RESULT   3
US-09-272-796-14
; Sequence 14, Application US/09272796
; Patent No. 6207148
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Guegler, Karl G.
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  DISEASE ASSOCIATED PROTEIN
;    TITLE OF INVENTION:  KINASES
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/272,796
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/878,989
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J J
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0321 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1435 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  PENITUT01
;      CLONE:  1452972
US-09-272-796-14


  Query Match             22.3%;  Score 366.6;  DB 4;  Length 1435;
  Best Local Similarity   62.9%;  Pred. No. 2.8e-87;
  Matches  567;  Conservative    0;  Mismatches  334;  Indels    0;  Gaps    0;

Qy     567 tctacatgcgcttcatgcaggagcacacctgctacgatgccatggcaactagctccaagc 626
           | |||||||| ||||||  |   ||||  || || ||   | |  |||| || || ||||
Db     183 TTTACATGCGATTCATGAGGTCACACAAGTGTTATGACATCGTTCCAACCAGTTCAAAGC 242

Qy     627 tagtcatcttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacg 686
           | ||  |||| || || |   |  | | || ||||||||||||||| |||| |||||||
Db     243 TTGTTGTCTTTGATACTACATTACAAGTTAAAAAGGCCTTCTTTGCTTTGGTAGCCAACG 302

Qy     687 gtgtgcgggcagcccctctatgggacagcaagaagcagagctttgtggggatgctgacca 746
           |||| || ||||| || || ||||| || || || || || ||||| || ||||| || |
Db     303 GTGTCCGAGCAGCGCCACTGTGGGAGAGTAAAAAACAAAGTTTTGTAGGAATGCTAACAA 362

Qy     747 tcactgacttcatcctggtgctgcatcgctactacaggtccccccctggtccagatctatg 806
           | || || |||||      | || ||| | |||||  |  || ||  |||| ||||| ||||
Db     363 TTACAGATTTCATAAATATACTACATAGATACTATAAATCACCTATGGTACAGATTTATG 422

Qy     807 agattgaacaacataagattgagacctggagggagatctacctgcaaggctgcttcaagc 866
           |  | ||  |||||||| ||||| || |||||||||| | ||  | ||||     || ||||
Db     423 AATTAGAGGAACATAAAATTGAAACATGGAGGGAGCTTTATTTACAAGAAACATTTAAGC 482

Qy     867 ctctggtctccatctctcctaatgatagcctgtttgaagctgtctacaccctcatcaaga 926
           || | ||     || |||||  |||  ||||| || || ||||| ||| || | ||||| |
Db     483 CTTTAGTGAATATATCTCCAGATGCAAGCCTCTTCGATGCTGTATACTCCTTGATCAAAA 542

Qy     927 accggatccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcctca 986
           |    |||||  |  |||| ||| |||||||  |    || || | |||  | || || |
Db     543 ATAAAATCCACAGATTGCCCGTTATTGACCCTATCAGTGGGAATGCACTTTATATACTTA 602

Qy     987 cacacaaacgcctgctcaagttcctgcacatctttggttccctgctgccccggccctcct 1046
           | |||||| |  | ||||||||||| ||  | |||   ||     ||||  |||  |||
Db     603 CCCACAAAAGAATCCTCAAGTTCCTCCAGCTTTTTATGTCTGATATGCCAAAGCCTGCCT 662

Qy    1047 tcctctaccgcactatccaagatttgggcatcggcacattccgagacttggctgtggtgc 1106
           || |  | |   |   | ||  | || || || || | ||   || | ||  |  | |
Db     663 TCATGAAGCAGAACCTGGATGAGCTTGGAATAGGAACGTACCACAACATTGCCTTCATAC 722

Qy    1107 tggagacagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctgcac 1166
                | |||||| | |  ||  || |||| ||||||||  | ||  | || || |
Db     723 ATCCAGACACTCCCATCATCAAAGCCTTGAACATATTTGTGGAAAGACGAATATCAGCTC 782

Qy    1167 tgcctgtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtgattc 1226
           |||||||  ||  | ||  |  ||  | || ||  |   | ||||||   |||||||| |||
Db     783 TGCCTGTTGTGGATGAGTCAGGAAAAGTTGTAGATATTTATTCCAAATTTGATGTAATTA 842

Qy    1227 acctggctgcccagcaaacctacaaccacctggacatgagtgtgggagaagccctgaggc 1286
           | || |||||  || |||| |||||  |||| || || |  |||    | |||||   ||
Db     843 ATCTTGCTGCTGAGAAAACATACAATAACCTAGATATCACGGTGACCCAGGCCCTTCAGC 902

Qy    1287 agaggacactatgtctggagggagtcctttcctgccagccccacgagagcttgggggaag 1346
           |  |  |||  | | | || || ||  |    ||| |    |  || |   ||| |
Db     903 ACCGTTCACAGTATTTTGAAGGTGTTGTGAAGTGCAATAAGCTGGAAATACTGGAGACCA 962
```

```
Qy   1347 tgatcgacaggattgctcgggagcaggtacacaggctggtgctagtggacgagacccagc 1406
          |  | ||||| || |   | |   |||| ||  |||||||| | ||  | ||  |  |
Db    963 TCGTGGACAGAATAGTAAGAGCTGAGGTCCATCGGCTGGTGGTGGTAAATGAAGCAGATA 1022

Qy   1407 atctcttgggcgtggtctccctctccgacatccttcaggcactggtgctcagccctgctg 1466
           | |  ||||  |  | ||||| || ||||| || || || ||| | ||||  || || |
Db   1023 GTATTGTGGGTATTATTTCCCTGTCGGACATTCTGCAAGCCCTGATCCTCACACCAGCAG 1082

Qy   1467 g 1467
           |
Db   1083 G 1083


RESULT   4
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14


  Query Match             3.0%;  Score 49;  DB 1;  Length 7218;
  Best Local Similarity   2.6%;  Pred. No. 0.0021;
  Matches   10;  Conservative  220;  Mismatches  155;  Indels    0;  Gaps    0;

Qy    864 agcctctggtctccatctctcctaatgatagcctgtttgaagctgtctacaccctcatca 923
          ||| |  | |:::: :::::::::  :  :  ::: :::     :: ::: : ::::: ::
Db   1057 AGCTTGCGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1116

Qy    924 agaaccggatccatcgcctgcctgttcttgacccggtgtcaggcaacgtactccacatcc 983
```

```
          ::   ::: :: ::: ::: :::::  :::  : ::   :  : : :::: : :::
Db   1117 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1176

Qy    984 tcacacacaaacgcctgctcaagttcctgcacatctttggttccctgctgccccggccct 1043
          :: : : :   : ::: :::   ::::: : : :::::  :::::: :: :::  ::::
Db   1177 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1236

Qy   1044 ccttcctctaccgcactatccaagatttgggcatcggcacattccgagacttggctgtgg 1103
          ::::::::: :: : :: :::    :::   : ::  : : ::::    :::  :: :
Db   1237 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1296

Qy   1104 tgctggagacagcacccatcctgactgcactggacatctttgtggaccggcgtgtgtctg 1163
          : ::    :  : ::: ::::  :: : ::   : ::::: :   ::  : : : :::
Db   1297 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1356

Qy   1164 cactgcctgtggtcaacgaatgtggtcaggtcgtgggcctctattcccgctttgatgtga 1223
          : :: ::: :  ::  :   : :  ::   :: :   ::::: ::::: ::::  : :
Db   1357 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1416

Qy   1224 ttcacctggctgcccagcaaaccta 1248
          ::: :::  :: :::  :  ||| |
Db   1417 YYYYYYYYYYYYYYYYYYGTACCAA 1441


RESULT   5
US-09-056-105-8
; Sequence 8, Application US/09056105
; Patent No. 6287569
; GENERAL INFORMATION:
;  APPLICANT: KIPPS, THOMAS J.
;  APPLICANT: WU, YUNQI
;  TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR
;  TITLE OF INVENTION: PROCESSING
;  FILE REFERENCE: 233/221
;  CURRENT APPLICATION NUMBER: US/09/056,105
;  CURRENT FILING DATE: 1998-04-06
;  EARLIER APPLICATION NUMBER: 60/043,467
;  EARLIER FILING DATE: 1997-04-10
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 8
;   LENGTH: 1022
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-056-105-8


  Query Match             2.5%;  Score 41.2;  DB 4;  Length 1022;
  Best Local Similarity   44.7%;  Pred. No. 0.12;
  Matches  160;  Conservative    0;  Mismatches  198;  Indels    0;  Gaps    0;

Qy    373 ctctgactgtacagcctcagctgcaggctccagcacagatgatgtggagctggccacgga 432
          | ||  ||| |       |   ||  || |||||||   | || | ||||  |||||
Db    299 cacttgctggaggcaacccaatgagggttccagcagccaagaagaggaggaggccagcac 358

Qy    433 gttcccagccacagaggcctgggagtgtgagctagaaggcctgctggaagagaggcctgc 492
          |  || | | ||||| ||| |     | ||   |    |       ||  ||   |  ||
Db    359 ctcgcctgacgcagagtccttgttccgagaagcactcagtaacaaggtggatgagttggc 418

Qy    493 cctgtgcctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaa 552
          |  |  |||  ||   ||   | |       ||||||  |    | |  ||| ||| | |
Db    419 tcatttttctgctccgcaagtatcgagccaaggagctggtcacaaaggcagaaatgctgga 478

Qy    553 acccggcgcccagatctacatgcgcttcatgcaggagcacacctgctacgatgccatggc 612
          |  |  | | |  |||| ||||| | | |  | |  |  | || | |   ||  | |
Db    479 gagagtcatcaaaaattacaagcgctgctttcctgtgatcttcggcaaagcctccgagtc 538

Qy    613 aactagctccaagctagtcatcttcgacaccatgctggagatcaagaaggccttctttgc 672
          |     |   | | |||   ||  |      ||||   |   |    |  ||   |
Db    539 cctgaagatgatctttggcattgacgtgaaggaagtggaccccgccagcaacacctacac 598

Qy    673 tctggtggccaacggtgtgcgggcagcccctctatgggacagcaagaagcagagcttt 730
          || ||  ||  |   |  |   |     |     |    | || || |||| ||||
```

```
Db    599 ccttgtcacctgcctgggccttcctatgatggcctgctgggtaataatcagatcttt 656


RESULT   6
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
;  GENERAL INFORMATION:
;    APPLICANT:  DORNER, F.
;    APPLICANT:  SCHEIFLINGER, F.
;    APPLICANT:  FALKNER, F. G.
;    TITLE OF INVENTION:  RECOMBINANT FOWLPOX VIRUS
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Foley & Lardner
;      STREET:  1800 Diagonal Road, Suite 500
;      CITY:  Alexandria
;      STATE:  VA
;      COUNTRY:  USA
;      ZIP:  22313-0299
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/232,463
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/935,313
;      FILING DATE:
;      APPLICATION NUMBER:  EP 91 114 300.6
;      FILING DATE:  26-AUG-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  BENT, Stephen A.
;      REGISTRATION NUMBER:  29,768
;      REFERENCE/DOCKET NUMBER:  30472/114 IMMU
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (703)836-9300
;      TELEFAX:  (703)683-4109
;      TELEX:  899149
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7218 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      CLONE:  pTZgpt-F1s
US-08-232-463-14


  Query Match             2.4%;  Score 40.2;  DB 1;  Length 7218;
  Best Local Similarity   4.5%;  Pred. No. 0.44;
  Matches   12;  Conservative  150;  Mismatches  103;  Indels    0;  Gaps    0;

Qy      8 ggggctggccacatggagcccgggctggagcacgcactgcgcaggacccccttcctggagc 67
          ::::  ::  : : ::::   :::  :::: : : :  : : ::::          ::::
Db   1307 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1248

Qy     68 agccttgggggttctgagcatcaagaqatgagcttcctagagcaagaaaacagcagctca 127
          ::    :::::    ::: :  :::::: :::       :::: ::::::: :: ::   :
Db   1247 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1188

Qy    128 tggccatcaccagctgtgaccagcagctcagaaagaatccgtgggaaacggagggccaaa 187
           ::  :  :  ::  : ::  :: ::   ::::::::   : :::::: ::::::  :::
Db   1187 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1128

Qy    188 gccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtcaggggaa 247
          :   :::: ::: :::: :::::  :: :::::::::::::  :  :::  ::::::::
Db   1127 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1068

Qy    248 ggtccccggtccaggccagctgctg 272
          : || |  |  |   | | |||| |
Db   1067 RATCGCAAGCTCCCTCGACCTGCAG 1043


RESULT   7
US-09-056-105-18
; Sequence 18, Application US/09056105
; Patent No. 6287569
; GENERAL INFORMATION:
;  APPLICANT: KIPPS, THOMAS J.
;  APPLICANT: WU, YUNQI
;  TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR
;  TITLE OF INVENTION: PROCESSING
;  FILE REFERENCE: 233/221
;  CURRENT APPLICATION NUMBER: US/09/056,105
;  CURRENT FILING DATE: 1998-04-06
;  EARLIER APPLICATION NUMBER: 60/043,467
;  EARLIER FILING DATE: 1997-04-10
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 18
;   LENGTH: 1022
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-056-105-18


  Query Match             2.4%;  Score 39.6;  DB 4;  Length 1022;
  Best Local Similarity   44.4%;  Pred. No. 0.31;
  Matches  159;  Conservative    0;  Mismatches  199;  Indels    0;  Gaps    0;

Qy    373 ctctgactgtacagcctcagctgcaggctccagcacagatgatgtggagctggccacgga 432
          | ||  ||| |       |   ||  || |||||||   | || | ||||  | | |
Db    299 cacttgctggaggcaacccaatgagggttccagcagccaagaagaggaggggccaagcac 358

Qy    433 gttcccagccacagaggcctgggagtgtgagctagaaggcctgctggaagagaggcctgc 492
           |  || | | ||||| ||| |    | ||   |     |       ||  ||   |   ||
Db    359 ctcgcctgacgcagagtccttgttccgagaagcactcagtaacaaggtggatgagttggc 418

Qy    493 cctgtgcctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaa 552
           |  |  |||  ||   ||   | |       ||||||  |    | |  ||| ||| | |
Db    419 tcattttctgctccgcaagtatcgagccaaggagctggtcacaaaggcagaaatgctgga 478

Qy    553 acccggcgcccagatctacatgcgcttcatgcaggagcacacctgctacgatgccatggc 612
             | |  | | |  |||| ||||| | | |  | |  |  | || | |   ||  | |
Db    479 gagagtcatcaaaaattacaagcgctgctttcctgtgatcttcggcaaagcctccgagtc 538

Qy    613 aactagctccaagctagtcatcttcgacaccatgctggagatcaagaaggccttctttgc 672
             |     |  | | |||   ||  |      ||||   ||  |    |  ||   |
Db    539 cctgaagatgatctttggcattgacgtgaaggaagtggaccccaccagcaacacctacac 598

Qy    673 tctggtggccaacggtgtgcgggcagcccctctatgggacagcaagaagcagagcttt 730
          || ||  ||  |   |  |  |      |     |    | || || |||| ||||
Db    599 ccttgtcacctgcctgggcctttcctatgatggcctgctgggtaataatcagatcttt 656


RESULT   8
US-08-299-849B-14
; Sequence 14, Application US/08299849B
; Patent No. 5612201
;  GENERAL INFORMATION:
;    APPLICANT:  De Plaen, Etienne; Boon-Falleur, Thierry;
;    APPLICANT:  Leth , Bernard; Szikora, Jean-Pierre; De Smet, Charles;
;    APPLICANT:  Chomez, Patrick
;    TITLE OF INVENTION:  Isolated Nucleic Acid Molecules Useful In
;    TITLE OF INVENTION:  Determining Expression Of A Tumor Antigen Precursor
;    NUMBER OF SEQUENCES:  48
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
```

```
;       ZIP:  10022
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
;       COMPUTER:  IBM
;       OPERATING SYSTEM:  PC-DOS
;       SOFTWARE:  Wordperfect
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/299,849B
;       FILING DATE:  1-SEPTEMBER-1994
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/037,230
;       FILING DATE:  26-MARCH-1993
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US92/04354
;       FILING DATE:  22-MAY-1992
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/807,043
;       FILING DATE:  12-DECEMBER-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/764,364
;       FILING DATE:  23-SEPTEMBER-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/728,838
;       APPLICATION NUMBER:  9-JULY-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/705,702
;       FILING DATE:  23-May-1991
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Hanson, No. 5612201man D.
;       REGISTRATION NUMBER:  30,946
;       REFERENCE/DOCKET NUMBER:  LUD 5355
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 688-9200
;       TELEFAX:  (212) 838-3884
;   INFORMATION FOR SEQ ID NO:  14:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2531 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  genomic DNA
;     FEATURE:
;       NAME/KEY:  MAGE-41 gene
US-08-299-849B-14


  Query Match             2.4%;  Score 39.6;  DB 1;  Length 2531;
  Best Local Similarity   44.4%;  Pred. No. 0.43;
  Matches  159;  Conservative    0;  Mismatches  199;  Indels    0;  Gaps    0;

Qy     373 ctctgactgtacagcctcagctgcaggctccagcacagatgatgtggagctggccacgga 432
           | ||  ||| |      |   ||  || |||||||   | || | ||||  | | |
Db     858 CACTTGCTGGAGGCAACCCAATGAGGGTTCCAGCAGCCAAGAAGAGGAGGGGCCAAGCAC 917

Qy     433 gttcccagccacagaggcctgggagtgtgagctagaaggcctgctggaagagaggcctgc 492
             | || | | ||||| ||| |    | ||   |    |       ||  ||   |  ||
Db     918 CTCGCCTGACGCAGAGTCCTTGTTCCGAGAAGCACTCAGTAACAAGGTGGATGAGTTGGC 977

Qy     493 cctgtgcctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaa 552
             |  |  |||  ||   ||   | |       ||||||  |    | |  ||| ||| | |
Db     978 TCATTTTCTGCTCCGCAAGTATCGAGCCAAGGAGCTGGTCACAAAGGCAGAAATGCTGGA 1037

Qy     553 acccggcgcccagatctacatgcgcttcatgcaggagcacacctgctacgatgccatggc 612
               | |  | | |  |||| ||||| | | |  | |  |  | || | |   ||  | |
Db    1038 GAGAGTCATCAAAAATTACAAGCGCTGCTTTCCTGTGATCTTCGGCAAAGCCTCCGAGTC 1097

Qy     613 aactagctccaagctagtcatcttcgacaccatgctggagatcaagaaggccttctttgc 672
              |      |  | | |||   ||  |      ||||   ||  |    |  ||   |
Db    1098 CCTGAAGATGATCTTTGGCATTGACGTGAAGGAAGTGGACCCCACCAGCAACACCTACAC 1157

Qy     673 tctggtggccaacggtgtgcgggcagcccctctatgggacagcaagaagcagagctttt 730
           || ||  ||  |   |  |  |   |     |    | || || |||| ||||
```

```
Db    1158 CCTTGTCACCTGCCTGGGCCTTTCCTATGATGGCCTGCTGGGTAATAATCAGATCTTT 1215


RESULT   9
US-08-142-368A-14
; Sequence 14, Application US/08142368A
; Patent No. 5925729
;  GENERAL INFORMATION:
;    APPLICANT:  Boon-Falleur, Thierry; Van der Bruggen, Thierry;
;    APPLICANT:  Van den Eynde, Beno t; Van Pel, Aline; De Plaen, Etienne;
;    APPLICANT:  Lurquin, Christophe; Chomez, Patrick; Traversari, Catia
;    TITLE OF INVENTION:  Tumor Rejection Antigen Precursors, Tumor
;    TITLE OF INVENTION:  Rejection Antigens and Uses Thereof
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
;      COMPUTER:  IBM
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/142,368A
;      FILING DATE:  02-MAY-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/04354
;      FILING DATE:  22-MAY-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/807,043
;      FILING DATE:  12-DECEMBER-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/764,364
;      FILING DATE:  23-SEPTEMBER-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/728,838
;      APPLICATION NUMBER:  9-JULY-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/705,702
;      FILING DATE:  23-May-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 5925729man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  LUD 5253.4-US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2531 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  genomic DNA
;    FEATURE:
;      NAME/KEY:  MAGE-41 gene
US-08-142-368A-14


  Query Match             2.4%;  Score 39.6;  DB 2;  Length 2531;
  Best Local Similarity   44.4%;  Pred. No. 0.43;
  Matches  159;  Conservative    0;  Mismatches  199;  Indels    0;  Gaps    0;

Qy     373 ctctgactgtacagcctcagctgcaggctccagcacagatgatgtggagctggccacgga 432
           | ||  ||| |      |   ||  || |||||||   | || | ||||  | | |
Db     858 CACTTGCTGGAGGCAACCCAATGAGGGTTCCAGCAGCCAAGAAGAGGAGGGGCCAAGCAC 917

Qy     433 gttcccagccacagaggcctgggagtgtgagctagaaggcctgctggaagagaggcctgc 492
             | || | | ||||| ||| |    | ||   |    |       ||  ||   |  ||
```

```
Db    918 CTCGCCTGACGCAGAGTCCTTGTTCCGAGAAGCACTCAGTAACAAGGTGGATGAGTTGGC 977

Qy    493 cctgtgcctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaa 552
          |  |  |||  ||   ||   | |       ||||||  |     | |  ||| ||| | |
Db    978 TCATTTTCTGCTCCGCAAGTATCGAGCCAAGGAGCTGGTCACAAAGGCAGAAATGCTGGA 1037

Qy    553 acccggcgcccagatctacatgcgcttcatgcaggagcacacctgctacgatgccatggc 612
          | |  | | |  |||| ||||| | | |  | |  |  | || | |   ||  | |
Db   1038 GAGAGTCATCAAAAATTACAAGCGCTGCTTTCCTGTGATCTTCGGCAAAGCCTCCGAGTC 1097

Qy    613 aactagctccaagctagtcatcttcgacaccatgctggagatcaagaaggccttctttgc 672
          |     |   | | |||   ||  |      ||||  ||  |    |  ||   |
Db   1098 CCTGAAGATGATCTTTGGCATTGACGTGAAGGAAGTGGACCCCACCAGCAACACCTACAC 1157

Qy    673 tctggtggccaacggtgtgcgggcagcccctctatgggacagcaagaagcagagcttt 730
          || ||  ||  |   |  |   |      |     |    | || || |||| ||||
Db   1158 CCTTGTCACCTGCCTGGGCCTTTCCTATGATGGCCTGCTGGGTAATAATCAGATCTTT 1215


RESULT 10
US-08-967-727-14
; Sequence 14, Application US/08967727
; Patent No. 6025474
;  GENERAL INFORMATION:
;    APPLICANT:  Gaugler, B atrice; Van den Eynde, Beno t;
;    APPLICANT:  van der Bruggen, Pierre;  Boon-Falleur, Thierry
;    TITLE OF INVENTION:  Isolated Nucleic Acid Molecules Coding For
;    TITLE OF INVENTION:  Tumor Rejection Antigen Precursor Mage-3 And Uses Thereof
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
;      COMPUTER:  IBM
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/967,727
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/037,230
;      FILING DATE:  26-MARCH-1993
;      APPLICATION NUMBER:  PCT/US92/04354
;      FILING DATE:  22-MAY-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/807,043
;      FILING DATE:  12-DECEMBER-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/764,365
;      FILING DATE:  23-SEPTEMBER-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/728,838
;      FILING DATE:  9-JULY-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/705,702
;      FILING DATE:  23-MAY-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 6025474man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  LUD 5353
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2531 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  genomic DNA
;    FEATURE:
;      NAME/KEY:  MAGE-41 gene
US-08-967-727-14


  Query Match             2.4%;  Score 39.6;  DB 3;  Length 2531;
  Best Local Similarity   44.4%;  Pred. No. 0.43;
  Matches  159;  Conservative    0;  Mismatches  199;  Indels    0;  Gaps    0;

Qy    373 ctctgactgtacagcctcagctgcaggctccagcacagatgatgtggagctggccacgga 432
          | ||  ||| |      |   ||  || |||||||  | || | ||||  | | |
Db    858 CACTTGCTGGAGGCAACCCAATGAGGGTTCCAGCAGCCAAGAAGAGGAGGGGCCAAGCAC 917

Qy    433 gttcccagccacagaggcctgggagtgtgagctagaaggcctgctggaagagaggcctgc 492
          |  || | | ||||| ||| |    | ||   |    |       ||  ||   |   ||
Db    918 CTCGCCTGACGCAGAGTCCTTGTTCCGAGAAGCACTCAGTAACAAGGTGGATGAGTTGGC 977

Qy    493 cctgtgcctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaa 552
          |  |  |||  ||   ||   | |       ||||||  |     | |  ||| ||| | |
Db    978 TCATTTTCTGCTCCGCAAGTATCGAGCCAAGGAGCTGGTCACAAAGGCAGAAATGCTGGA 1037

Qy    553 acccggcgcccagatctacatgcgcttcatgcaggagcacacctgctacgatgccatggc 612
          | |  | | |  |||| ||||| | | |  | |  |  | || | |   ||  | |
Db   1038 GAGAGTCATCAAAAATTACAAGCGCTGCTTTCCTGTGATCTTCGGCAAAGCCTCCGAGTC 1097

Qy    613 aactagctccaagctagtcatcttcgacaccatgctggagatcaagaaggccttctttgc 672
          |     |   | | |||   ||  |      ||||  ||  |    |  ||   |
Db   1098 CCTGAAGATGATCTTTGGCATTGACGTGAAGGAAGTGGACCCCACCAGCAACACCTACAC 1157

Qy    673 tctggtggccaacggtgtgcgggcagcccctctatgggacagcaagaagcagagcttt 730
          || ||  ||  |   |  |   |      |     |    | || || |||| ||||
Db   1158 CCTTGTCACCTGCCTGGGCCTTTCCTATGATGGCCTGCTGGGTAATAATCAGATCTTT 1215


RESULT 11
US-08-037-230D-14
; Sequence 14, Application US/08037230D
; Patent No. 6235525
;  GENERAL INFORMATION:
;    APPLICANT:  Gaugler, B atrice; Van den Eynde, Beno t;
;    APPLICANT:  van der Bruggen, Pierre;  Boon-Falleur, Thierry
;    TITLE OF INVENTION:  Isolated Nucleic Acid Molecules Coding For
;    TITLE OF INVENTION:  Tumor Rejection Antigen Precursor Mage-3 And Uses Thereof
;    NUMBER OF SEQUENCES:  30
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Felfe & Lynch
;      STREET:  805 Third Avenue
;      CITY:  New York City
;      STATE:  New York
;      ZIP:  10022
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 360 kb storage
;      COMPUTER:  IBM
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  Wordperfect
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/037,230D
;      FILING DATE:  26-MARCH-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US92/04354
;      FILING DATE:  22-MAY-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/807,043
;      FILING DATE:  12-DECEMBER-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/764,364
;      FILING DATE:  23-SEPTEMBER-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/728,838
;      FILING DATE:  9-JULY-1991
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/705,702
;      FILING DATE:  23-MAY-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hanson, No. 6235525man D.
;      REGISTRATION NUMBER:  30,946
;      REFERENCE/DOCKET NUMBER:  LUD 5353
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 688-9200
;      TELEFAX:  (212) 838-3884
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2531 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  genomic DNA
;    FEATURE:
;      NAME/KEY:  MAGE-41 gene
US-08-037-230D-14


  Query Match             2.4%;  Score 39.6;  DB 4;  Length 2531;
  Best Local Similarity   44.4%;  Pred. No. 0.43;
  Matches  159;  Conservative    0;  Mismatches  199;  Indels    0;  Gaps    0;

Qy     373 ctctgactgtacagcctcagctgcaggctccagcacagatgatgtggagctggccacgga 432
           | ||  ||| |      |   ||  || |||||||   | || | ||||  | | |
Db     858 CACTTGCTGGAGGCAACCCAATGAGGGTTCCAGCAGCCAAGAAGAGGAGGGGCCAAGCAC 917

Qy     433 gttcccagccacagaggcctgggagtgtgagctagaaggcctgctggaagagaggcctgc 492
           |  || | | ||||| ||| |    | ||   |    |      ||  ||   |   ||
Db     918 CTCGCCTGACGCAGAGTCCTTGTTCCGAGAAGCACTCAGTAACAAGGTGGATGAGTTGGC 977

Qy     493 cctgtgcctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaa 552
           |  |  |||  ||   ||   | |       ||||||  |    | |  ||| ||| | |
Db     978 TCATTTTCTGCTCCGCAAGTATCGAGCCAAGGAGCTGGTCACAAAGGCAGAAATGCTGGA 1037

Qy     553 acccggcgcccagatctacatgcgcttcatgcaggagcacacctgctacgatgccatggc 612
           |  |  | | |  |||| ||||| | | |  | |  |  | || | |   ||  | |
Db    1038 GAGAGTCATCAAAAATTACAAGCGCTGCTTTCCTGTGATCTTCGGCAAAGCCTCCGAGTC 1097

Qy     613 aactagctccaagctagtcatcttcgacaccatgctggagatcaagaaggccttctttgc 672
           |     |   | | |||   ||  |      ||||   ||  |    |  ||   |
Db    1098 CCTGAAGATGATCTTTGGCATTGACGTGAAGGAAGTGGACCCCACCAGCAACACCTACAC 1157

Qy     673 tctggtggccaacggtgtgcgggcagcccctctatgggacagcaagaagcagagcttt 730
           || ||  ||  |   |  |   |      |     | || || |||| ||||
Db    1158 CCTTGTCACCTGCCTGGGCCTTTCCTATGATGGCCTGCTGGGTAATAATCAGATCTTT 1215


RESULT  12
US-09-056-105-10
; Sequence 10, Application US/09056105
; Patent No. 6287569
; GENERAL INFORMATION:
;  APPLICANT: KIPPS, THOMAS J.
;  APPLICANT: WU, YUNQI
;  TITLE OF INVENTION: VACCINES WITH ENHANCED INTRACELLULAR
;  TITLE OF INVENTION: PROCESSING
;  FILE REFERENCE: 233/221
;  CURRENT APPLICATION NUMBER: US/09/056,105
;  CURRENT FILING DATE: 1998-04-06
;  EARLIER APPLICATION NUMBER: 60/043,467
;  EARLIER FILING DATE: 1997-04-10
;  NUMBER OF SEQ ID NOS: 35
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 10
;   LENGTH: 4895
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-056-105-10


  Query Match             2.4%;  Score 39.6;  DB 4;  Length 4895;
  Best Local Similarity   44.4%;  Pred. No. 0.55;
  Matches  159;  Conservative    0;  Mismatches  199;  Indels    0;  Gaps    0;

Qy     373 ctctgactgtacagcctcagctgcaggctccagcacagatgatgtggagctggccacgga 432
           | ||  ||| |      |   ||  || |||||||   | || | ||||  | | |
Db    3222 cacttgctggaggcaacccaatgagggttccagcagccaagaagaggaggggccaagcac 3281

Qy     433 gttcccagccacagaggcctgggagtgtgagctagaaggcctgctggaagagaggcctgc 492
           |  || | | ||||| ||| |    | ||   |    |      ||  ||   |   ||
Db    3282 ctcgcctgacgcagagtccttgttccgagaagcactcagtaacaaggtggatgagttggc 3341

Qy     493 cctgtgcctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaa 552
           |  |  |||  ||   ||   | |       ||||||  |    | |  ||| ||| | |
Db    3342 tcattttctgctccgcaagtatcgagccaaggagctggtcacaaaggcagaaatgctgga 3401

Qy     553 acccggcgcccagatctacatgcgcttcatgcaggagcacacctgctacgatgccatggc 612
           |  |  | | |  |||| ||||| | | |  | |  |  | || | |   ||  | |
Db    3402 gagagtcatcaaaaattacaagcgctgctttcctgtgatcttcggcaaagcctccgagtc 3461

Qy     613 aactagctccaagctagtcatcttcgacaccatgctggagatcaagaaggccttctttgc 672
           |     |   | | |||   ||  |      ||||   ||  |    |  ||   |
Db    3462 cctgaagatgatctttggcattgacgtgaaggaagtggaccccaccagcaacacctacac 3521

Qy     673 tctggtggccaacggtgtgcgggcagcccctctatgggacagcaagaagcagagcttt 730
           || ||  ||  |   |  |   |      |     | || || |||| ||||
Db    3522 ccttgtcacctgcctgggcctttcctatgatggcctgctgggtaataatcagatcttt 3579


RESULT  13
US-09-007-005-17/c
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
;  APPLICANT: Szostak, Jack W.
;  APPLICANT: Roberts, Richard W.
;  APPLICANT: Liu, Rihe
;  TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
;  TITLE OF INVENTION: FUSIONS
;  FILE REFERENCE: 00786/350003
;  CURRENT APPLICATION NUMBER: US/09/007,005B
;  CURRENT FILING DATE: 1998-01-14
;  EARLIER APPLICATION NUMBER: 60/035,963
;  EARLIER FILING DATE: 1997-01-27
;  EARLIER APPLICATION NUMBER: 60/064,491
;  EARLIER FILING DATE: 1997-11-06
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 289
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Translation template
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(289)
;   OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17


  Query Match             2.4%;  Score 38.8;  DB 4;  Length 289;
  Best Local Similarity   5.3%;  Pred. No. 0.31;
  Matches   10;  Conservative   87;  Mismatches   93;  Indels    0;  Gaps    0;

Qy     894 gcctgtttgaagctgtctacaccctcatcaagaaccggatccatcgcctgcctgttcttg 953
           |: :|:|:  | |:|:|:|: :|:|: ::     ::   ::: :  ::   :   : :
Db     246 GYAYGYTYTYAYCYGYCYAYGYCYTYGYSYNYNYSYNYNYSYNYNYSYNYNYSY 187

Qy     954 acccggtgtcaggcaacgtactccacatcctcacacacaaacgcctgctcaagttcctgc 1013
           : ::    :   :  :    ::: : ::: : : : :   : : ::: :  :: : :::
Db     186 NYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSY 127
```

```
Qy   1014 acatctttggttccctgctgccccggccctccttcctctaccgcactatccaaagatttgg 1073
          : ::: :   : : :::   :::   ::: : ::: :  :: : :   :     : ::
Db    126 NYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSY 67

Qy   1074 gcatcggcac 1083
          : ::  : :
Db     66 NYNYSYNYNY 57


RESULT 14
US-09-244-796-17/c
; Sequence 17, Application US/09244796
; Patent No. 6281344
; GENERAL INFORMATION:
;  APPLICANT: Szostak, Jack W.
;  APPLICANT: Roberts, Richard W.
;  APPLICANT: Liu, Rihe
;  TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
;  TITLE OF INVENTION: FUSIONS
;  FILE REFERENCE: 00786/350007
;  CURRENT APPLICATION NUMBER: US/09/244,796
;  CURRENT FILING DATE: 1999-02-05
;  EARLIER APPLICATION NUMBER: 60/035,963
;  EARLIER FILING DATE: 1997-01-27
;  EARLIER APPLICATION NUMBER: 60/064,491
;  EARLIER FILING DATE: 1997-11-06
;  EARLIER APPLICATION NUMBER: 09/007,005
;  EARLIER FILING DATE: 1998-01-14
;  NUMBER OF SEQ ID NOS: 33
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
;   LENGTH: 289
;   TYPE: RNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Translation template
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(289)
;   OTHER INFORMATION: n = A,T,C or G
US-09-244-796-17


  Query Match             2.4%;  Score 38.8;  DB 4;  Length 289;
  Best Local Similarity   5.3%;  Pred. No. 0.31;
  Matches   10;  Conservative   87;  Mismatches   93;  Indels    0;  Gaps    0;

Qy    894 gcctgtttgaagctgtctacaccctcatcaagaaccggatccatcgcctgcctgttcttg 953
          |: :|:|:  | |:|:|:|: :|:|: ::     ::   ::: :  ::   :   : :
Db    246 GYAYGYTYTYAYCYGYCYAYGYCYTYGYSYNYNYSYNYNYSYNYNYSYNYNYSY 187

Qy    954 acccggtgtcaggcaacgtactccacatcctcacacacaaacgcctgctcaagttcctgc 1013
          : ::    :  :  :   ::: : ::: : : : :   : : ::: :  :: : :::
Db    186 NYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSY 127

Qy   1014 acatctttggttccctgctgccccggccctccttcctctaccgcactatccaaagatttgg 1073
          : ::: :   : : :::   :::   ::: : ::: :  :: : :   :     : ::
Db    126 NYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSYNYNYSY 67

Qy   1074 gcatcggcac 1083
          : ::  : :
Db     66 NYNYSYNYNY 57


RESULT 15
US-09-404-650-1
; Sequence 1, Application US/09404650
; Patent No. 6309858
; GENERAL INFORMATION:
;  APPLICANT: Dietrich, Paul S.
;  APPLICANT: McGivern, Joseph G.
;  TITLE OF INVENTION: T-TYPE CALCIUM CHANNEL VARIANTS; COMPOSITIONS THEREOF;
;  TITLE OF INVENTION: AND USES
;  FILE REFERENCE: R0043B-REG sequence listing
;  CURRENT APPLICATION NUMBER: US/09/404,650
;  CURRENT FILING DATE: 1999-09-23
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 6816
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (192)..(6716)
US-09-404-650-1


  Query Match             2.4%;  Score 38.8;  DB 4;  Length 6816;
  Best Local Similarity   47.5%;  Pred. No. 1;
  Matches  115;  Conservative    0;  Mismatches  127;  Indels    0;  Gaps    0;

Qy    631 catcttcgacaccatgctggagatcaagaaggccttctttgctctggtggccaacggtgt 690
          || ||||||||| | ||    ||      || | || ||||||  ||||   |||| | ||
Db   4169 caacttcgacaacctgggccaggctctgatgtccctctttgtcctggcatccaaggatgg 4228

Qy    691 gcgggcagccccctctatgggacagcaagaagcagagctttgtggggatgctgaccatcac 750
           |||    |    | |   |  |   | |          |||||    || | |  | ||
Db   4229 ttgggtgaacatcatgtacaatggactggatgctgttgctgtggaccagcagcctgtgac 4288

Qy    751 tgacttcatcctggtgctgcatcgctactacaggtcccccctggtccagatctatgagat 810
            ||  || ||    | |||  |  |||| ||  |||  |||| ||    ||       |
Db   4289 caaccacaacccctggatgctgctgtacttcatctccttcctgctcatcgtcagcttctt 4348

Qy    811 tgaacaacataagattgagacctggagggagatctacctgcaaggctgcttcaagcctct 870
          ||  |   | | | ||| |         || |    || | ||    |||    |||  |
Db   4349 tgtgctcaacatgtttgtgggtgtcgtggtggagaacttccacaagtgccggcagcacca 4408

Qy    871 gg 872
          ||
Db   4409 gg 4410


Search completed: October  3, 2002, 16:22:08
Job time: 16968 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 14:48:55 ; Search time 7316.32 Seconds
(without alignments)
288.886 Million cell updates/sec

Title: US-09-826-581-5_COPY_180_280
Perfect score: 101
Sequence: 1 gggccaaagccttgagatgg.....ggccagctgctgagtccacc 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB	ID	Description
1	101	100.0	1647	6	AX281582	AX281582 Sequence
2	101	100.0	2290	9	HSA249977	AJ249977 Homo sapi
3	99.4	98.4	2109	6	AX099776	AX099776 Sequence
4	99.4	98.4	2115	6	AX099802	AX099802 Sequence
5	99.4	98.4	2115	9	AF214519	AF214519 Homo sapi
6	72.2	71.5	1867	6	AX099774	AX099774 Sequence
7	72.2	71.5	1873	4	AF214520	AF214520 Sus scrof
8	72.2	71.5	1873	6	AX099800	AX099800 Sequence
9	72.2	71.5	2022	6	AX099804	AX099804 Sequence
10	71	70.3	989	6	AX281579	AX281579 Sequence
c 11	71	70.3	152129	2	AC027416	AC027416 Homo sapi
c 12	71	70.3	206854	9	AC009974	AC009974 Homo sapi
c 13	69.4	68.7	196554	2	AC073128	AC073128 Homo sapi
14	45.4	45.0	5888	4	AF214521	AF214521 Sus scrof
c 15	42.2	41.8	227724	2	AF336381	AF336381 Mus muscu
16	33	32.7	155012	8	AP003257	AP003257 Oryza sat
17	32	31.7	160133	2	AC024355	AC024355 Homo sapi
18	32	31.7	160878	2	AC016971	AC016971 Homo sapi
19	32	31.7	162840	2	AC046203	AC046203 Homo sapi
c 20	32	31.7	191330	9	AC025271	AC025271 Homo sapi
21	32	31.7	192175	2	AL591490	AL591490 Mus muscu
c 22	31.8	31.5	8670	8	CRARG7	X16619 Chlamydomon
c 23	31.6	31.3	124057	9	AP000770	AP000770 Homo sapi
c 24	31.6	31.3	129750	2	AC103965	AC103965 Homo sapi
c 25	31.6	31.3	153201	9	AC006454	AC006454 Homo sapi
c 26	31.6	31.3	165655	2	AC027605	AC027605 Homo sapi
c 27	31.6	31.3	166941	2	AC073891	AC073891 Homo sapi
c 28	31.6	31.3	194101	2	AC090910	AC090910 Homo sapi
c 29	31.6	31.3	210691	9	AP001891	AP001891 Homo sapi
c 30	31.4	31.1	121851	2	AC091751	AC091751 Gallus ga
31	31.2	30.9	126380	2	AP000714	AP000714 Homo sapi
32	31.2	30.9	139505	9	HSJ591C20	AL118506 Human DNA
33	31.2	30.9	142746	2	AC093564	AC093564 Homo sapi
c 34	31.2	30.9	157473	2	AL356977	AL356977 Homo sapi
c 35	31.2	30.9	159840	2	AP000869	AP000869 Homo sapi
36	31.2	30.9	171902	2	AP000846	AP000846 Homo sapi
37	31.2	30.9	198410	2	AP000831	AP000831 Homo sapi
38	31.2	30.9	199321	9	AP000941	AP000941 Homo sapi
39	31	30.7	31549	9	AC096857	AC096857 Homo sapi
40	31	30.7	78505	2	AC006408	AC006408 Homo sapi
41	31	30.7	123551	2	AC005809	AC005809 Homo sapi
c 42	31	30.7	191832	2	AC091723	AC091723 Sus scrof
43	30.8	30.5	7459	1	SC8A2	AL445327 Streptomy
c 44	30.8	30.5	36583	3	LMFL2464	AL365154 Leishmani
c 45	30.8	30.5	177648	3	LMFP265	AL359716 Leishmani

ALIGNMENTS

RESULT 1
AX281582
LOCUS AX281582 1647 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 5 from Patent WO0177305.
ACCESSION AX281582
VERSION AX281582.1 GI:16608833
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Andersson,L., Luthman,H. and Marklund,S.
TITLE Variants of the human amp-activated protein kinase gamma 3 subunit
JOURNAL Patent: WO 0177305-A 5 18-OCT-2001;
Arexis AB (SE)
FEATURES Location/Qualifiers
source 1. .1647
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS 20. .1489
/note="unnamed protein product"
/codon_start=1

```
                     /protein_id="CAD10589.1"
                     /db_xref="GI:16608834"
                     /translation="MEPGLEHALRRTPSWSSLGGSEHQEMSFLEQENSSSWPSPAVTS
                     SSERIRGKRRAKALRWTRQKSVEEGEPPGQGEGPRSRPAAESTGLEATFPKTTPLAQA
                     DPAGVGTPPTGWDCLPSDCTASAAGSSTDDVELATEFPATEAWECELEGLLEERPALC
                     LSPQAPFPKLGWDDELRKPGAQIYMRFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFA
                     LVANGVRAAPLWDSKKQSFVGMLTITDFILVLHRYYRSPLVQIYEIEQHKIETWREIY
                     LQGCFKPLVSISPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIF
                     GSLLPRPSFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSALPVVNECGQV
                     VGLYSRFDVIHLAAQQTYNHLDMSVGEALRQRTLCLEGVLSCQPHESLGEVIDRIARE
                     QVHRLVLVDETQHLLGVVSLSDILQALVLSPAGIDALGA"
BASE COUNT      346 a    502 c    462 g    337 t
ORIGIN


  Query Match           100.0%;  Score 101;  DB 6;  Length 1647;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-16;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 GGGCCAAAGCCTTGAGATGGACAAGGCAGAAGTCGGTGGAGGAAGGGGAGCCACCAGGTC 239

Qy     61 agggggaaggtccccggtccaggccagctgctgagtccacc 101
          |||||||||||||||||||||||||||||||||||||||||
Db    240 AGGGGGAAGGTCCCCGGTCCAGGCCAGCTGCTGAGTCCACC 280


RESULT   2
HSA249977
LOCUS       HSA249977               2290 bp    mRNA    linear   PRI 07-APR-2000
DEFINITION  Homo sapiens mRNA for AMP-activated protein kinase gamma 3 subunit
            (AMPK gamma 3 gene).
ACCESSION   AJ249977
VERSION     AJ249977.1  GI:6688200
KEYWORDS    AMP-activated protein kinase; AMPK gamma 3 gene; gamma 3 subunit.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2290)
  AUTHORS   Cheung,P.C., Salt,I.P., Davies,S.P., Hardie,D.G. and Carling,D.
  TITLE     Characterization of AMP-activated protein kinase gamma-subunit
            isoforms and their role in AMP binding
  JOURNAL   Biochem. J. 346 Pt 3, 659-669 (2000)
  MEDLINE   20164049
REFERENCE   2  (bases 1 to 2290)
  AUTHORS   Carling,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-OCT-1999) Carling D., Cellular Stress Group, MRC
            Clinical Sciences Centre, Hammersmith Hospital, DuCane Road,
            London, W12 0NN, UNITED KINGDOM
FEATURES             Location/Qualifiers
     source          1. .2290
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            22. .1500
                     /gene="AMPK gamma 3"
     CDS             22. .1500
                     /gene="AMPK gamma 3"
                     /function="AMP-activated protein kinase regulatory
                     subunit"
                     /codon_start=1
                     /evidence=experimental
                     /product="AMP-activated protein kinase gamma 3 subunit"
                     /protein_id="CAB65117.1"
                     /db_xref="GI:6688201"
                     /translation="MEPGLEHALRRTPSWSSLGGSEHQEMSFLEQENSSSWPSPAVTS
                     SSERIRGKRRAKALRWTRQKSVEEGEPPGQGEGPRSRPAAESTGLEATFPKTTPLAQA
                     DPAGVGTPPTGWDCLPSDCTASAAGSSTDDVELATEFPATEAWECELEGLLEERPALC
                     LSPQAPFPKLGWDDELRKPGAQIYMRFIEEHTCYDAMATSSKLVIFDTMLEIKKAFFA
                     LVANGVRAAPLWDSKKQSFVGMLTITDFILVLHRYYRSPLVQIYEIEQHKIETWREIY
                     LQGCFKPLVSISPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIF
                     GSLLPRPSFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSALPVVNECGQV
                     VGLYSRFDVIHLAAQQTYNHLDMSVGEALRKRTLCLEGVLSCQPHESLGEVIDRIARE
                     QVHRLVLVDETQHLLGVVSLSDILQALVLSPAGIDPSGPEKI"
BASE COUNT      501 a    674 c    617 g    498 t
ORIGIN


  Query Match           100.0%;  Score 101;  DB 9;  Length 2290;
  Best Local Similarity  100.0%;  Pred. No. 1.6e-16;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    182 GGGCCAAAGCCTTGAGATGGACAAGGCAGAAGTCGGTGGAGGAAGGGGAGCCACCAGGTC 241

Qy     61 agggggaaggtccccggtccaggccagctgctgagtccacc 101
          |||||||||||||||||||||||||||||||||||||||||
Db    242 AGGGGGAAGGTCCCCGGTCCAGGCCAGCTGCTGAGTCCACC 282


RESULT   3
AX099776
LOCUS       AX099776                2109 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 3 from Patent WO0120003.
ACCESSION   AX099776
VERSION     AX099776.1  GI:13538810
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2109)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 3 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .2109
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             472. .1389
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35799.1"
                     /db_xref="GI:13538811"
                     /translation="MRFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAP
                     LWDSKKQSFVGMLTITDFILVLHRYYRSPLVQIYEIEQHKIETWREIYLQGCFKPLVS
                     ISPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLPRPSFL
                     YRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSALPVVNECGQVVGLYSRFDVI
                     HLAAQQTYNHLDMSVGEALRQRTLCLEGVLSCQPHESLGEVIDRIAREQVHRLVLVDE
                     TQHLLGVVSLSDILQALVLSPAGIDALGA"
BASE COUNT      458 a    621 c    560 g    470 t
ORIGIN


  Query Match           98.4%;  Score 99.4;  DB 6;  Length 2109;
  Best Local Similarity  99.0%;  Pred. No. 4.1e-16;
  Matches  100;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     80 GGGCCAAAGCCTTGAGATGGACAAGGCAGAAGTCGGTGGAGGAAGGGGAGCCACCAGGTC 139

Qy     61 agggggaaggtccccggtccaggccagctgctgagtccacc 101
          |||||||||||||||||||||||| ||||||||||||||
Db    140 AGGGGGAAGGTCCCCGGTCCAGGCCAACTGCTGAGTCCACC 180


RESULT   4
```

```
AX099802
LOCUS       AX099802               2115 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 29 from Patent WO0120003.
ACCESSION   AX099802
VERSION     AX099802.1  GI:13538836
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2115)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 29 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .2115
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             1. .1395
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35801.1"
                     /db_xref="GI:13538837"
                     /translation="MSFLEQENSSSWPSPAVTSSSERIRGKRRAKALRWTRQKSVEEG
                     EPPGQGEGPRSRPTAESTGLEATFPKTTPLAQADPAGVGTPPTGWDCLPSDCTASAAG
                     SSTDDVELATEFPATEAWECELEGLLEERPALCLSPQAPFPKLGWDDELRKPGAQIYM
                     RFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEQHKIETWREIYLQGCFKPLVSISPNDSLFEAVYTLI
                     KNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNECGQVVGLYSRFDVIHLAAQQTYNHLDMSV
                     GEALRQRTLCLEGVLSCQPHESLGEVIDRIAREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT      460 a    622 c    562 g    471 t
ORIGIN
```

```
  Query Match             98.4%;  Score 99.4;  DB 6;  Length 2115;
  Best Local Similarity   99.0%;  Pred. No. 4.1e-16;
  Matches  100;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86 GGGCCAAAGCCTTGAGATGGACAAGGCAGAAGTCGGTGGAGGAAGGGGAGCCACCAGGTC 145

Qy      61 agggggaaggtccccggtccaggccagctgctgagtccacc 101
           ||||||||||||||||||||||||||| ||||||||||||||
Db     146 AGGGGGAAGGTCCCCGGTCCAGGCCAACTGCTGAGTCCACC 186
```

```
RESULT   5
AF214519
LOCUS       AF214519               2115 bp    mRNA    linear   PRI 03-JUN-2000
DEFINITION  Homo sapiens AMP-activated protein kinase gamma subunit (PRKAG3)
            mRNA, complete cds.
ACCESSION   AF214519
VERSION     AF214519.1  GI:8215681
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2115)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A., Thelander,M.,
            Rogel-Gaillard,C., Paul,S., Iannuccelli,N., Rask,L., Ronne,H.,
            Lundstrom,K., Reinsch,N., Gellin,J., Kalm,E., Roy,P.L., Chardon,P.
            and Andersson,L.
  TITLE     A mutation in PRKAG3 associated with excess glycogen content in pig
            skeletal muscle
  JOURNAL   Science 288 (5469), 1248-1251 (2000)
  MEDLINE   20280150
   PUBMED   10818001
REFERENCE   2  (bases 1 to 2115)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A.,
            Rogel-Gaillard,C., Paul,S., Gellin,J., Lundstrom,K., Reinsch,N.,
            Kalm,E., Le Roy,P., Chardon,P. and Andersson,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-1999) Dept Animal Breeding and Genetics, Swedish
            University of Agricultural Sciences, BMC box 597, Uppsala 751 24,
            Sweden
FEATURES             Location/Qualifiers
     source          1. .2115
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2p"
                     /tissue_type="skeletal muscle"
     gene            1. .2115
                     /gene="PRKAG3"
     CDS             1. .1395
                     /gene="PRKAG3"
                     /note="AMPKG3"
                     /codon_start=1
                     /product="AMP-activated protein kinase gamma subunit"
                     /protein_id="AAF73987.1"
                     /db_xref="GI:8215682"
                     /translation="MSFLEQENSSSWPSPAVTSSSERIRGKRRAKALRWTRQKSVEEG
                     EPPGQGEGPRSRPTAESTGLEATFPKTTPLAQADPAGVGTPPTGWDCLPSDCTASAAG
                     SSTDDVELATEFPATEAWECELEGLLEERPALCLSPQAPFPKLGWDDELRKPGAQIYM
                     RFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEQHKIETWREIYLQGCFKPLVSISPNDSLFEAVYTLI
                     KNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNECGQVVGLYSRFDVIHLAAQQTYNHLDMSV
                     GEALRQRTLCLEGVLSCQPHESLGEVIDRIAREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT      460 a    622 c    562 g    471 t
ORIGIN
```

```
  Query Match             98.4%;  Score 99.4;  DB 9;  Length 2115;
  Best Local Similarity   99.0%;  Pred. No. 4.1e-16;
  Matches  100;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86 GGGCCAAAGCCTTGAGATGGACAAGGCAGAAGTCGGTGGAGGAAGGGGAGCCACCAGGTC 145

Qy      61 agggggaaggtccccggtccaggccagctgctgagtccacc 101
           ||||||||||||||||||||||||||| ||||||||||||||
Db     146 AGGGGGAAGGTCCCCGGTCCAGGCCAACTGCTGAGTCCACC 186
```

```
RESULT   6
AX099774
LOCUS       AX099774               1867 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 1 from Patent WO0120003.
ACCESSION   AX099774
VERSION     AX099774.1  GI:13538808
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 1867)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 1 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
```

```
FEATURES             Location/Qualifiers
     source          1. .1867
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
     CDS             472. .1389
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35798.1"
                     /db_xref="GI:13538809"
                     /translation="MHFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAP
                     LWDSKKQSFVGMLTITDFILVLHRYYRSPLVQIYEIEEHKIETWREIYLQGCFKPLVS
                     ISPNDSLFEAVYALIKNRIHRLPVLDPVSGAVLHILTHKRLLKFLHIFGTLLPRPSFL
                     YRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSALPVVNETGQVVGLYSRFDVI
                     HLAAQQTYNHLDMNVGEALRQRTLCLEGVLSCQPHETLGEVIDRIVREQVHRLVLVDE
                     TQHLLGVVSLSDILQALVLSPAGIDALGA"
BASE COUNT      380 a    583 c    529 g    375 t
ORIGIN
```

```
  Query Match           71.5%;  Score 72.2;  DB 6;  Length 1867;
  Best Local Similarity  82.2%;  Pred. No. 4.6e-09;
  Matches   83;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           ||  ||| ||||  ||||||||||||||| ||   || |||||||||| ||| || || |
Db      80 GGAACAAGGCCTCTAGATGGACAAGGCAGGAGGATGTAGAGGAAGGGGGGCCTCCGGGCC 139

Qy      61 agggggaaggtccccggtccaggccagctgctgagtccacc 101
             | |||||||||||| |||||||||| ||||||||||||
Db     140 CGAGGGAAGGTCCCCAGTCCAGGCCAGTTGCTGAGTCCACC 180
```

```
RESULT   7
AF214520
LOCUS       AF214520               1873 bp    mRNA    linear   MAM 03-JUN-2000
DEFINITION  Sus scrofa AMP-activated protein kinase gamma subunit (PRKAG3)
            mRNA, complete cds.
ACCESSION   AF214520
VERSION     AF214520.1  GI:8215683
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 1873)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A., Thelander,M.,
            Rogel-Gaillard,C., Paul,S., Iannuccelli,N., Rask,L., Ronne,H.,
            Lundstrom,K., Reinsch,N., Gellin,J., Kalm,E., Roy,P.L., Chardon,P.
            and Andersson,L.
  TITLE     A mutation in PRKAG3 associated with excess glycogen content in pig
            skeletal muscle
  JOURNAL   Science 288 (5469), 1248-1251 (2000)
  MEDLINE   20280150
   PUBMED   10818001
REFERENCE   2  (bases 1 to 1873)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A.,
            Rogel-Gaillard,C., Paul,S., Gellin,J., Lundstrom,K., Reinsch,N.,
            Kalm,E., Le Roy,P., Chardon,P. and Andersson,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-1999) Dept Animal Breeding and Genetics, Swedish
            University of Agricultural Sciences, BMC box 597, Uppsala 751 24,
            Sweden
FEATURES             Location/Qualifiers
     source          1. .1873
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /chromosome="15"
                     /map="15q"
                     /tissue_type="skeletal muscle"
     gene            1. .1873
                     /gene="PRKAG3"
     CDS             1. .1395
                     /gene="PRKAG3"
                     /note="AMPKG3"
                     /codon_start=1
                     /product="AMP-activated protein kinase gamma subunit"
                     /protein_id="AAF73988.1"
                     /db_xref="GI:8215684"
                     /translation="MSFLEQGESRSWPSRAVTTSSERSHGDQGNKASRWTRQEDVEEG
                     GPPGPREGPQSRPVAESTGQEATFPKATPLAQAAPLAEVDNPPTERDILPSDCAASAS
                     DSNTDHLDLGIEFSASAASGDELGLVEEKPAPCPSPEVLLPRLGWDDELQKPGAQVYM
                     HFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEEHKIETWREIYLQGCFKPLVSISPNDSLFEAVYALI
                     KNRIHRLPVLDPVSGAVLHILTHKRLLKFLHIFGTLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNETGQVVGLYSRFDVIHLAAQQTYNHLDMNV
                     GEALRQRTLCLEGVLSCQPHETLGEVIDRIVREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT      382 a    580 c    535 g    376 t
ORIGIN
```

```
  Query Match           71.5%;  Score 72.2;  DB 4;  Length 1873;
  Best Local Similarity  82.2%;  Pred. No. 4.6e-09;
  Matches   83;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           ||  ||| ||||  ||||||||||||||| ||   || |||||||||| ||| || || |
Db      86 GGAACAAGGCCTCTAGATGGACAAGGCAGGAGGATGTAGAGGAAGGGGGGCCTCCGGGCC 145

Qy      61 agggggaaggtccccggtccaggccagctgctgagtccacc 101
             | |||||||||||| |||||||||| ||||||||||||
Db     146 CGAGGGAAGGTCCCCAGTCCAGGCCAGTTGCTGAGTCCACC 186
```

```
RESULT   8
AX099800
LOCUS       AX099800               1873 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 27 from Patent WO0120003.
ACCESSION   AX099800
VERSION     AX099800.1  GI:13538834
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 1873)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 27 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .1873
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
     CDS             1. .1395
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35800.1"
                     /db_xref="GI:13538835"
                     /translation="MSFLEQGESRSWPSRAVTTSSERSHGDQGNKASRWTRQEDVEEG
                     GPPGPREGPQSRPVAESTGQEATFPKATPLAQAAPLAEVDNPPTERDILPSDCAASAS
                     DSNTDHLDLGIEFSASAASGDELGLVEEKPAPCPSPEVLLPRLGWDDELQKPGAQVYM
                     HFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEEHKIETWREIYLQGCFKPLVSISPNDSLFEAVYALI
                     KNRIHRLPVLDPVSGAVLHILTHKRLLKFLHIFGTLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNETGQVVGLYSRFDVIHLAAQQTYNHLDMNV
                     GEALRQRTLCLEGVLSCQPHETLGEVIDRIVREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT      382 a    580 c    535 g    376 t
ORIGIN
```

```
  Query Match             71.5%;  Score 72.2;  DB 6;  Length 1873;
  Best Local Similarity   82.2%;  Pred. No. 4.6e-09;
  Matches   83;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           ||  ||| ||||  ||||||||||||||| ||   || |||||||||| ||| || || |
Db      86 GGAACAAGGCCTCTAGATGGACAAGGCAGGAGGATGTAGAGGAAGGGGGGCCTCCGGGCC 145

Qy      61 agggggaaggtccccggtccaggccagctgctgagtccacc 101
             | ||||||||||| |||||||||| ||||||||||||
Db     146 CGAGGGAAGGTCCCCAGTCCAGGCCAGTTGCTGAGTCCACC 186


RESULT   9
AX099804
LOCUS       AX099804               2022 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 31 from Patent WO0120003.
ACCESSION   AX099804
VERSION     AX099804.1  GI:13538838
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 2022)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 31 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .2022
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
BASE COUNT      412 a    623 c    593 g    394 t
ORIGIN


  Query Match             71.5%;  Score 72.2;  DB 6;  Length 2022;
  Best Local Similarity   82.2%;  Pred. No. 4.5e-09;
  Matches   83;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           ||  ||| ||||  ||||||||||||||| ||   || |||||||||| ||| || || |
Db     236 GGAACAAGGCCTCTAGATGGACAAGGCAGGAGGATGTAGAGGAAGGGGGGCCTCCGGGCC 295

Qy      61 agggggaaggtccccggtccaggccagctgctgagtccacc 101
             | ||||||||||| |||||||||| ||||||||||||
Db     296 CGAGGGAAGGTCCCCAGTCCAGGCCAGTTGCTGAGTCCACC 336


RESULT  10
AX281579
LOCUS       AX281579               989 bp    DNA     linear   PAT 02-NOV-2001
DEFINITION  Sequence 2 from Patent WO0177305.
ACCESSION   AX281579
VERSION     AX281579.1  GI:16608830
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Andersson,L., Luthman,H. and Marklund,S.
  TITLE     Variants of the human amp-activated protein kinase gamma 3 subunit
  JOURNAL   Patent: WO 0177305-A 2 18-OCT-2001;
            Arexis AB (SE)
FEATURES             Location/Qualifiers
     source          1. .989
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      229 a    306 c    286 g    168 t
ORIGIN


  Query Match             70.3%;  Score 71;  DB 6;  Length 989;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-08;
  Matches   71;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     109 GGGCCAAAGCCTTGAGATGGACAAGGCAGAAGTCGGTGGAGGAAGGGGAGCCACCAGGTC 168

Qy      61 agggggaaggt 71
           |||||||||||
Db     169 AGGGGGAAGGT 179


RESULT  11
AC027416/c
LOCUS       AC027416               152129 bp    DNA     linear   HTG 07-JUN-2000
DEFINITION  Homo sapiens clone RP11-504G11, WORKING DRAFT SEQUENCE, 32
            unordered pieces.
ACCESSION   AC027416
VERSION     AC027416.2  GI:8317289
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 152129)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-504G11
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 152129)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jun 7, 2000 this sequence version replaced gi:7342115.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L7458
                Center clone name: 504_G_11
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
```

```
    Chemistry: Dye-terminator Big Dye; 100% of reads
    Assembly program: Phrap; version 0.960731
    Consensus quality: 135376 bases at least Q40
    Consensus quality: 143264 bases at least Q30
    Consensus quality: 146503 bases at least Q20
    Insert size: 161000; agarose-fp
    Insert size: 149029; sum-of-contigs
    Quality coverage: 3.1 in Q20 bases; agarose-fp
    Quality coverage: 3.3 in Q20 bases; sum-of-contigs
--------------.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 32 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1     1005: contig of 1005 bp in length
*     1006 1105: gap of      100 bp
*     1106     2402: contig of 1297 bp in length
*     2403 2502: gap of      100 bp
*     2503     3823: contig of 1321 bp in length
*     3824 3923: gap of      100 bp
*     3924     5020: contig of 1097 bp in length
*     5021 5120: gap of      100 bp
*     5121     6161: contig of 1041 bp in length
*     6162 6261: gap of      100 bp
*     6262     7547: contig of 1286 bp in length
*     7548 7647: gap of      100 bp
*     7648     9983: contig of 2336 bp in length
*     9984 10083: gap of      100 bp
*    10084    12556: contig of 2473 bp in length
*    12557 12656: gap of      100 bp
*    12657    15043: contig of 2387 bp in length
*    15044 15143: gap of      100 bp
*    15144    17123: contig of 1980 bp in length
*    17124 17223: gap of      100 bp
*    17224    19466: contig of 2243 bp in length
*    19467 19566: gap of      100 bp
*    19567    21928: contig of 2362 bp in length
*    21929 22028: gap of      100 bp
*    22029    24319: contig of 2291 bp in length
*    24320 24419: gap of      100 bp
*    24420    27059: contig of 2640 bp in length
*    27060 27159: gap of      100 bp
*    27160    30170: contig of 3011 bp in length
*    30171 30270: gap of      100 bp
*    30271    33968: contig of 3698 bp in length
*    33969 34068: gap of      100 bp
*    34069    38179: contig of 4111 bp in length
*    38180 38279: gap of      100 bp
*    38280    42366: contig of 4087 bp in length
*    42367 42466: gap of      100 bp
*    42467    46365: contig of 3899 bp in length
*    46366 46465: gap of      100 bp
*    46466    51285: contig of 4820 bp in length
*    51286 51385: gap of      100 bp
*    51386    55871: contig of 4486 bp in length
*    55872 55971: gap of      100 bp
*    55972    60595: contig of 4624 bp in length
*    60596 60695: gap of      100 bp
*    60696    66595: contig of 5900 bp in length
*    66596 66695: gap of      100 bp
*    66696    73218: contig of 6523 bp in length
*    73219 73318: gap of      100 bp
*    73319    77115: contig of 3797 bp in length
*    77116 77215: gap of      100 bp
*    77216    85022: contig of 7807 bp in length
*    85023 85122: gap of      100 bp
*    85123    93314: contig of 8192 bp in length
*    93315 93414: gap of      100 bp
*    93415   101193: contig of 7779 bp in length
*   101194 101293: gap of      100 bp
*   101294   113090: contig of 11797 bp in length
*   113091 113190: gap of      100 bp
*   113191   123496: contig of 10306 bp in length
*   123497 123596: gap of      100 bp
*   123597   137837: contig of 14241 bp in length
*   137838 137937: gap of      100 bp
*   137938   152129: contig of 14192 bp in length.
FEATURES             Location/Qualifiers
     source          1. .152129
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-504G11"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1005
                     /note="assembly_fragment"
     misc_feature    1106. .2402
                     /note="assembly_fragment"
     misc_feature    2503. .3823
                     /note="assembly_fragment"
     misc_feature    3924. .5020
                     /note="assembly_fragment"
     misc_feature    5121. .6161
                     /note="assembly_fragment"
     misc_feature    6262. .7547
                     /note="assembly_fragment"
     misc_feature    7648. .9983
                     /note="assembly_fragment"
     misc_feature    10084. .12556
                     /note="assembly_fragment"
     misc_feature    12657. .15043
                     /note="assembly_fragment"
     misc_feature    15144. .17123
                     /note="assembly_fragment"
     misc_feature    17224. .19466
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    19567. .21928
                     /note="assembly_fragment"
     misc_feature    22029. .24319
                     /note="assembly_fragment"
     misc_feature    24420. .27059
                     /note="assembly_fragment"
     misc_feature    27160. .30170
                     /note="assembly_fragment"
     misc_feature    30271. .33968
                     /note="assembly_fragment"
     misc_feature    34069. .38179
                     /note="assembly_fragment"
     misc_feature    38280. .42366
                     /note="assembly_fragment"
     misc_feature    42467. .46365
                     /note="assembly_fragment"
     misc_feature    46466. .51285
                     /note="assembly_fragment"
     misc_feature    51386. .55871
                     /note="assembly_fragment"
     misc_feature    55972. .60595
                     /note="assembly_fragment"
     misc_feature    60696. .66595
                     /note="assembly_fragment"
     misc_feature    66696. .73218
                     /note="assembly_fragment"
     misc_feature    73319. .77115
                     /note="assembly_fragment"
     misc_feature    77216. .85022
                     /note="assembly_fragment"
     misc_feature    85123. .93314
                     /note="assembly_fragment"
     misc_feature    93415. .101193
                     /note="assembly_fragment"
     misc_feature    101294. .113090
```

```
                  /note="assembly_fragment"

  Query Match             70.3%;  Score 71;  DB 2;  Length 152129;
  Best Local Similarity   100.0%;  Pred. No. 3.8e-09;
  Matches   71;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2945 GGGCCAAAGCCTTGAGATGGACAAGGCAGAAGTCGGTGGAGGAAGGGGAGCCACCAGGTC 2886

Qy      61 agggqgaaggt 71
           ||||||||||||
Db    2885 AGGGGGAAGGT 2875


RESULT  12
AC009974/c
LOCUS       AC009974               206854 bp    DNA     linear   PRI 09-JAN-2002
DEFINITION  Homo sapiens BAC clone RP11-459I19 from 2, complete sequence.
ACCESSION   AC009974
VERSION     AC009974.9  GI:16799058
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 206854)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
REFERENCE   2  (bases 1 to 206854)
  AUTHORS   Harris,A. and Cotton,M.
  TITLE     The sequence of Homo sapiens BAC clone RP11-459I19
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 206854)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 206854)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-NOV-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 206854)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JAN-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   6  (bases 1 to 206854)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JAN-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Nov 8, 2001 this sequence version replaced gi:13431203.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0459I19
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. John D.
            McPherson, Department of Genetics, Washington University, St. Louis
            MO.  For additional information about the map position of this
            sequence, see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries.  Genomics 51:1-8.  The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at the Roswell Park Cancer Institute
            (http://bacpac.med.buffalo.edu)
            VECTOR:  pBACe3.6
            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is RP11-1077K22; the clone
            sequenced to the right is RP11-647O5.  Actual start of this clone
            is at base position 1 of RP11-459I19; actual end is at base
            position 206854 of RP11-459I19.

            Data from AC079810 and AC073128 was used to finish this clone,
            AC009974.  Polymorphisms have been identified between AC073128 and
            AC009974.  A single plasmid region exists beteen 38812-38903.  An
            unresolved tandem in the HERV SVA exists between 184390-185163.
            PCR suggests that approximately 1700 bps are missing.
FEATURES             Location/Qualifiers
     source          1. .206854
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2"
                     /clone="RP11-459I19"
                     /clone_lib="RPCI-11"
     misc_feature    1. .37
                     /note="match to EST BE314060 (NID:g9134719)"
     misc_feature    1. .37
                     /note="match to EST BI059713 (NID:g14467240)"
     misc_feature    1. .37
                     /note="match to EST BF183086 (NID:g11061273)"
     misc_feature    1. .37
                     /note="match to EST AL567345 (NID:g12920610)"
     misc_feature    1. .37
                     /note="match to EST AW880850 (NID:g8042860)"
     misc_feature    1. .37
                     /note="match to EST BF304755 (NID:g11251653)"
     misc_feature    1. .37
                     /note="similar to Homo sapiens EST BI114348
                     (NID:g14565249)"
     misc_feature    1. .37
                     /note="match to EST BG477625 (NID:g13409904)"
     misc_feature    1. .37
                     /note="match to EST BE047599 (NID:g8364652) tz39c01.y1"
     misc_feature    3. .37
                     /note="match to EST BE908408 (NID:g10402954)"
     misc_feature    4. .37
                     /note="match to EST AI670836 (NID:g4850567) wa04g10.x1"
     misc_feature    164. .662
                     /note="match to EST C05773 (NID:g1502549)"
     misc_feature    220. .221
                     /note="match to EST BG470047 (NID:g13402322)"
     misc_feature    281. .344
                     /note="match to EST BE047599 (NID:g8364652) tz39c01.y1"
     misc_feature    281. .344
```

```
                     /note="match to EST BE908408 (NID:g10402954)"
     misc_feature    281. .344
                     /note="match to EST BF304755 (NID:g11251653)"
     misc_feature    281. .344
                     /note="match to EST BF183086 (NID:g11061273)"
     misc_feature    281. .344
                     /note="similar to Homo sapiens EST BI114348
                     (NID:g14565249)"
     misc_feature    281. .344
                     /note="match to EST BG477625 (NID:g13409904)"
     misc_feature    281. .344
                     /note="match to EST BI059713 (NID:g14467240)"
     misc_feature    281. .344
                     /note="match to EST AL567345 (NID:g12920610)"
     misc_feature    281. .344
                     /note="match to EST BE314060 (NID:g9134719)"
     misc_feature    281. .344
                     /note="match to EST AI670836 (NID:g4850567) wa04g10.x1"
     misc_feature    281. .344
                     /note="match to EST AW880850 (NID:g8042860)"
     misc_feature    294. .344
                     /note="match to EST BG470047 (NID:g13402322)"
     misc_feature    complement(557. .558)
                     /note="match to EST AA043371 (NID:g1521226) zk53e10.r1"
     misc_feature    588. .929
                     /note="match to EST T64073 (NID:g667938) yc05d12.r1"
     misc_feature    594. .763
                     /note="match to EST BF304755 (NID:g11251653)"
     misc_feature    594. .763
                     /note="match to EST BG470047 (NID:g13402322)"
     misc_feature    594. .763
                     /note="match to EST AW880850 (NID:g8042860)"
     misc_feature    594. .731
                     /note="match to EST BE314060 (NID:g9134719)"
     misc_feature    594. .764
                     /note="match to EST BI059713 (NID:g14467240)"
     misc_feature    594. .763
                     /note="match to EST BG477625 (NID:g13409904)"
     misc_feature    594. .742
                     /note="match to EST BE047599 (NID:g8364652) tz39c01.y1"
     misc_feature    594. .763
                     /note="match to EST BE908408 (NID:g10402954)"
     misc_feature    594. .763
                     /note="match to EST AA043371 (NID:g1521226) zk53e10.r1"
     misc_feature    594. .763
                     /note="match to EST AI670836 (NID:g4850567) wa04g10.x1"
     misc_feature    594. .763
                     /note="match to EST BF183086 (NID:g11061273)"
     misc_feature    594. .763
                     /note="similar to Homo sapiens EST BI114348
                     (NID:g14565249)"
     misc_feature    599. .763
                     /note="match to EST AL567345 (NID:g12920610)"
     misc_feature    617. .1084
                     /note="match to EST AA481361 (NID:g2210913) zv44e01.r1"
     misc_feature    622. .763
                     /note="match to EST AI860958 (NID:g5514574) wl56f05.x1"
     misc_feature    622. .763
                     /note="similar to Mus musculus EST AI196847 (NID:g3749453)
                     ui67e01.x1"
     misc_feature    684. .763
                     /note="match to EST BG992568 (NID:g14396638)"
     misc_feature    962. .1084
                     /note="match to EST AI656812 (NID:g4740791) tt54b06.x1"
     misc_feature    967. .1084
                     /note="match to EST BE908408 (NID:g10402954)"
     misc_feature    967. .1085
                     /note="match to EST BF304755 (NID:g11251653)"
     misc_feature    967. .1091
                     /note="match to EST AA043371 (NID:g1521226) zk53e10.r1"
     misc_feature    967. .1071

Query Match          70.3%;  Score 71;  DB 9;  Length 206854;
  Best Local Similarity   100.0%;  Pred. No. 3.6e-09;
  Matches   71;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168897 GGGCCAAAGCCTTGAGATGGACAAGGCAGAAGTCGGTGGAGGAAGGGGAGCCACCAGGTC 168838

Qy      61 agggggaaggt 71
           ||||||||||||
Db 168837 AGGGGGAAGGT 168827


RESULT  13
AC073128/c
LOCUS       AC073128              196554 bp    DNA     linear   HTG 21-FEB-2001
DEFINITION  Homo sapiens chromosome 2 clone RP11-647O5, WORKING DRAFT SEQUENCE,
            17 unordered pieces.
ACCESSION   AC073128
VERSION     AC073128.3  GI:13027579
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 196554)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 196554)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JUN-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Feb 21, 2001 this sequence version replaced gi:8469048.

            ---------------- Genome Center ----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: H_NH0647O05
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 98%
            Sequencing vector: plasmid; 0%
            Chemistry: Dye-primer ET; 98% of reads
            Chemistry: Dye-terminator Big Dye; 0% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 187795 bases at least Q40
            Consensus quality: 190513 bases at least Q30
            Consensus quality: 192099 bases at least Q20
            Insert size: 200000; agarose-fp
            Insert size: 194954; sum-of-contigs
            Quality coverage: 5.58 in Q20 bases; agarose-fp
            Quality coverage: 5.67 in Q20 bases; sum-of-contigs
            -------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 17 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1157: contig of 1157 bp in length
            *     1158     1257: gap of unknown length
            *     1258     3600: contig of 2343 bp in length
            *     3601     3700: gap of unknown length
            *     3701     5103: contig of 1403 bp in length
            *     5104     5203: gap of unknown length
            *     5204     8524: contig of 3321 bp in length
            *     8525     8624: gap of unknown length
            *     8625    11856: contig of 3232 bp in length
```

```
            *    11857    11956: gap of unknown length
            *    11957    15783: contig of 3827 bp in length
            *    15784    15883: gap of unknown length
            *    15884    21906: contig of 6023 bp in length
            *    21907    22006: gap of unknown length
            *    22007    28887: contig of 6881 bp in length
            *    28888    28987: gap of unknown length
            *    28988    35255: contig of 6268 bp in length
            *    35256    35355: gap of unknown length
            *    35356    44642: contig of 9287 bp in length
            *    44643    44742: gap of unknown length
            *    44743    58275: contig of 13533 bp in length
            *    58276    58375: gap of unknown length
            *    58376    73816: contig of 15441 bp in length
            *    73817    73916: gap of unknown length
            *    73917    92140: contig of 18224 bp in length
            *    92141    92240: gap of unknown length
            *    92241   113337: contig of 21097 bp in length
            *   113338   113437: gap of unknown length
            *   113438   130325: contig of 16888 bp in length
            *   130326   130425: gap of unknown length
            *   130426   149287: contig of 18862 bp in length
            *   149288   149387: gap of unknown length
            *   149388   196554: contig of 47167 bp in length.
FEATURES             Location/Qualifiers
     source          1. .196554
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /clone="RP11-647O5"
     misc_feature    1. .1157
                     /note="assembly_name:Contig17"
     misc_feature    1258. .3600
                     /note="assembly_name:Contig18"
     misc_feature    3701. .5103
                     /note="assembly_name:Contig19
                     clone_end:T7
                     vector_side:right"
     misc_feature    5204. .8524
                     /note="assembly_name:Contig20"
     misc_feature    8625. .11856
                     /note="assembly_name:Contig21"
     misc_feature    11957. .15783
                     /note="assembly_name:Contig22"
     misc_feature    15884. .21906
                     /note="assembly_name:Contig23"
     misc_feature    22007. .28887
                     /note="assembly_name:Contig24"
     misc_feature    28988. .35255
                     /note="assembly_name:Contig25"
     misc_feature    35356. .44642
                     /note="assembly_name:Contig26"
     misc_feature    44743. .58275
                     /note="assembly_name:Contig27"
     misc_feature    58376. .73816
                     /note="assembly_name:Contig28"
     misc_feature    73917. .92140
                     /note="assembly_name:Contig29"
     misc_feature    92241. .113337
                     /note="assembly_name:Contig30"
     misc_feature    113438. .130325
                     /note="assembly_name:Contig31"
     misc_feature    130426. .149287
                     /note="assembly_name:Contig32"
     misc_feature    149388. .196554
                     /note="assembly_name:Contig33
                     clone_end:SP6
                     vector_side:right"
BASE COUNT    52296 a  46993 c  45889 g  49770 t   1606 others
ORIGIN

  Query Match             68.7%;  Score 69.4;  DB 2;  Length 196554;
  Best Local Similarity   98.6%;  Pred. No. 9.4e-09;
  Matches   70;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db   63318 GGGCCAAAGCCTTGAGATGGACAAGGCAGAAGTCGGTGGAGGAAGGGGAGGCACCAGGTC 63259

Qy      61 agggggaaggt 71
           ||||||||||||
Db   63258 AGGGGGAAGGT 63248


RESULT  14
AF214521
LOCUS       AF214521                5888 bp    DNA     linear   MAM 03-JUN-2000
DEFINITION  Sus scrofa AMPK gamma subunit (PRKAG3) gene, complete cds.
ACCESSION   AF214521
VERSION     AF214521.1  GI:8215685
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 5888)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A., Thelander,M.,
            Rogel-Gaillard,C., Paul,S., Iannuccelli,N., Rask,L., Ronne,H.,
            Lundstrom,K., Reinsch,N., Gellin,J., Kalm,E., Roy,P.L., Chardon,P.
            and Andersson,L.
  TITLE     A mutation in PRKAG3 associated with excess glycogen content in pig
            skeletal muscle
  JOURNAL   Science 288 (5469), 1248-1251 (2000)
  MEDLINE   20280150
   PUBMED   10818001
REFERENCE   2  (bases 1 to 5888)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A.,
            Rogel-Gaillard,C., Paul,S., Gellin,J., Lundstrom,K., Reinsch,N.,
            Kalm,E., Le Roy,P., Chardon,P. and Andersson,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-1999) Dept Animal Breeding and Genetics, Swedish
            University of Agricultural Sciences, BMC Box 597, Uppsala 751 24,
            Sweden
FEATURES             Location/Qualifiers
     source          1. .5888
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /chromosome="15"
                     /map="15q"
     mRNA            join(<1. .154,515. .918,1809. .1890,2349. .2407,2509. .2554,
                     2771. .2825,3027. .3153,3286. .3451,4578. .4615,4791. .4937,
                     5294. .>5410)
                     /gene="PRKAG3"
                     /product="AMPK gamma subunit"
     gene            <1. .>5410
                     /gene="PRKAG3"
     CDS             join(1. .154,515. .918,1809. .1890,2349. .2407,2509. .2554,
                     2771. .2825,3027. .3153,3286. .3451,4578. .4615,4791. .4937,
                     5294. .5410)
                     /gene="PRKAG3"
                     /codon_start=1
                     /product="AMPK gamma subunit"
                     /protein_id="AAF73989.1"
                     /db_xref="GI:8215686"
                     /translation="MSFLEQGESRSWPSRAVTTSSERSHGDQGNKASRWTRQEDVEEG
                     GPPGPREGPQSRPVAESTGQEATFPKATPLAQAAPLAEVDNPPTERDILPSDCAASAS
                     DSNTDHLDLGIEFSASAASGDELGLVEEKPAPCPSPEVLLPRLGWDDELQKPGAQVYM
                     HFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEEHKIETWREIYLQGCFKPLVSISPNDSLFEAVYALI
                     KNRIHRLPVLDPVSGAVLHILTHKRLLKFLHIFGTLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNETGQVVGLYSRFDVIHLAAQQTYNHLDMNV
                     GEALRQRTLCLEGVLSCQPHETLGEVIDRIVREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT    1274 a   1634 c   1638 g   1339 t      3 others
ORIGIN
```

```
  Query Match             45.0%;  Score 45.4;  DB 4;  Length 5888;
  Best Local Similarity   77.5%;  Pred. No. 0.032;
  Matches   55;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           ||  ||| ||||  ||||||||||||||| ||   || |||||||||| ||| || || |
Db      86 GGAACAAGGCCTCTAGATGGACAAGGCAGGAGGATGTAGAGGAAGGGGGGCCTCCGGGCC 145

Qy      61 agggggaaggt 71
             | ||||||||
Db     146 CGAGGGAAGGT 156


RESULT  15
AF336381/c
LOCUS       AF336381               227724 bp    DNA     linear   HTG 02-APR-2001
DEFINITION  Mus musculus chromosome 1 clone PAC510; PAC457, *** SEQUENCING IN
            PROGRESS ***, 3 unordered pieces.
ACCESSION   AF336381
VERSION     AF336381.1  GI:13507298
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 227724)
  AUTHORS   Rump,A., Hayes,C., Brown,S.D.M. and Rosenthal,A.
  TITLE     Mouse chromosome 1 genomic sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 227724)
  AUTHORS   Rump,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JAN-2001) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenberstr. 11, Jena 07745, Germany
COMMENT     * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    17869: contig of 17869 bp in length
            *    17870    17969: gap of unknown length
            *    17970    32746: contig of 14777 bp in length
            *    32747    32846: gap of unknown length
            *    32847   227724: contig of 194878 bp in length.
FEATURES             Location/Qualifiers
     source          1. .227724
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="1"
                     /clone="PAC510; PAC457"
BASE COUNT    57663 a  55423 c  56238 g  58199 t    201 others
ORIGIN


  Query Match             41.8%;  Score 42.2;  DB 2;  Length 227724;
  Best Local Similarity   74.6%;  Pred. No. 0.1;
  Matches   53;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           | | ||| || |  |||||||| || ||| || | || |||||||  || |||||||||
Db    8850 GAGTCAAGGCTTCCAGATGGACGAGACAGGAGGCCGTAGAGGAAGCAGAACCACCAGGTT 8791

Qy      61 agggggaaggt 71
             ||| ||||||
Db    8790 TGGGAGAAGGT 8780
```

```
Search completed: October  3, 2002, 14:49:24
Job time: 12084 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 16:30:41 ; Search time 719.93 Seconds
(without alignments)
240.868 Million cell updates/sec

Title: US-09-826-581-5_COPY_180_280
Perfect score: 101
Sequence: 1 gggccaaagccttgagatgg.....ggccagctgctgagtccacc 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	101	100.0	1647	22	AAH43685	PRKAG3 cDNA. Homo
	2	99.4	98.4	2109	22	AAD03296	Human AMPK gamma s
	3	99.4	98.4	2115	22	AAD03320	Human AMPK gamma s
	4	72.2	71.5	1867	22	AAD03295	Pig AMPK gamma sub
	5	72.2	71.5	1873	22	AAD03319	Pig AMPK gamma sub
	6	72.2	71.5	2022	22	AAD03321	Sus scrofa PRKAG3
	7	71	70.3	989	22	AAH43682	PRKAG3 intron 2 -
c	8	31.8	31.5	8670	20	AAZ32059	Human METH2 relate
c	9	31.8	31.5	8670	22	AAC90316	X16619 cDNA clone.
	10	30	29.7	711	23	AAS84420	DNA encoding novel
	11	29.8	29.5	2541	23	AAS84523	DNA encoding novel
	12	29.8	29.5	4125	24	AAD22004	Human transporters
c	13	29.4	29.1	886	22	AAH06841	Human cDNA clone (
c	14	29.4	29.1	1457	22	AAH14551	Human cDNA sequenc
c	15	29.4	29.1	1482	19	AAV59752	Human secreted pro
c	16	29.2	28.9	930	22	AAK52299	Human polynucleoti
c	17	29.2	28.9	935	22	AAK53283	Human polynucleoti
c	18	29	28.7	3210	23	ABL12861	Drosophila melanog
	19	29	28.7	5981	23	ABL12860	Drosophila melanog
c	20	29	28.7	14738	22	AAL02949	Human reproductive
c	21	29	28.7	14738	22	AAL02951	Human reproductive
c	22	28.8	28.5	5505	20	AAX83492	Rat T-type voltage
	23	28.6	28.3	268	17	AAT10719	5'-end of human be
	24	28.6	28.3	268	19	AAV41784	Target sequence us
	25	28.6	28.3	428	16	AAQ84934	Human geta-globin
	26	28.6	28.3	596	16	AAQ83707	Portion of human b
c	27	28.6	28.3	798	22	AAH67920	C glutamicum codin
c	28	28.6	28.3	921	22	AAF72062	Corynebacterium gl
c	29	28.6	28.3	942	21	AAC79994	Human secreted pro
	30	28.6	28.3	3935	21	AAA65169	Human beta/gamma-g
	31	28.6	28.3	6060	18	AAT48454	Human beta-globin
	32	28.6	28.3	349980	22	AAH68533	C glutamicum codin
c	33	28.4	28.1	2463	22	AAH16987	Human cDNA sequenc
c	34	28.2	27.9	1653	21	AAC58992	Human secreted pro
	35	28	27.7	423	22	AAS43330	DNA encoding novel
c	36	28	27.7	452	22	AAS43405	DNA encoding novel
c	37	28	27.7	452	22	AAI84199	Human polynucleoti
c	38	28	27.7	31813	22	ABA08137	Human ovarian and
c	39	28	27.7	31813	22	AAL06997	Human reproductive
c	40	27.8	27.5	1182	21	AAA27105	Human h-TRAAK cDNA
c	41	27.8	27.5	1182	22	AAH78636	Human mechanically
c	42	27.8	27.5	1218	21	AAA27106	Human h-TRAAK cDNA
c	43	27.8	27.5	1257	24	AAH99922	Nucleotide sequenc
c	44	27.8	27.5	1337	22	AAH78646	Human mechanically
c	45	27.8	27.5	1408	24	AAH99921	Nucleotide sequenc

ALIGNMENTS

RESULT 1
AAH43685
ID AAH43685 standard; cDNA; 1647 BP.
XX
AC AAH43685;
XX
DT 21-JAN-2002 (first entry)
XX
DE PRKAG3 cDNA.
XX
KW Human; AMP-activated protein kinase gamma 3 subunit; PRKAG3; variant;
KW metabolic disease; diabetes; obesity; substitution; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 20..1489
FT /*tag= a
FT /product= "PRKAG3"
FT variation 230
FT /*tag= b
FT /label= "C230G"
FT /note= "Causes P71A"
FT variation 559
FT /*tag= c
FT /label= "T559C"
FT /note= "Silent variation"
FT variation 1037
FT /*tag= d
FT /label= "C1037T"
FT /note= "Causes R340W"
XX

```
PN   WO200177305-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-SE00765.
XX
PR   07-APR-2000; 2000US-195665P.
XX
PA   (AREX-) AREXIS AB.
XX
PI   Andersson L,  Luthman H,  Marklund S;
XX
DR   WPI; 2001-657170/75.
DR   P-PSDB; QQB47679.
XX
PT   New variants of human AMP-activated protein kinase gamma3 subunit
PT   associated with a metabolic disease e.g. diabetes or obesity and method
PT   for determining a risk estimate of diseases in subject by detecting the
PT   variant -
XX
PS   Disclosure; Fig 5; 25pp; English.
XX
CC   This sequence represents the full length cDNA encoding the human
CC   AMP-activated protein kinase gamma 3 subunit (PRKAG3). Detecting
CC   the presence of the PRKAG3 DNA, or a variant, is useful in determining
CC   a risk estimate of a metabolic disease, such as diabetes or obesity,
CC   in a subject.  The variation may occur in exons 3, 4 or 10.  In exon
CC   3 variation may be a substitution of a G for a C at nucleotide 320,
CC   resulting in the amino acid substitution P71A; in exon 4 variation may
CC   be a substitution of a T for a C at nucleotide 550; and in exon 10
CC   variation may be a substitution of a T for a C at nucleotide 1037,
CC   resulting in the amino acid substitution R340W.  There may also be
CC   nucleotide variation in intron 6.  The numbering of these
CC   variations is based on the full length cDNA as given, rather than on
CC   position 1 of the open reading frame.
XX
SQ   Sequence 1647 BP; 346 A; 502 C; 462 G; 337 T; 0 other;


  Query Match             100.0%;  Score 101;  DB 22;  Length 1647;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-20;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 239

Qy      61 agggggaaggtccccggtccaggccagctgctgagtccacc 101
           |||||||||||||||||||||||||||||||||||||||||
Db     240 agggggaaggtccccggtccaggccagctgctgagtccacc 280


RESULT   2
AAD03296
ID   AAD03296 standard; DNA; 2109 BP.
XX
AC   AAD03296;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Human AMPK gamma subunit muscle-specific isoform, PRKAG3 cDNA.
XX
KW   Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW   PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW   genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW   cystathione beta synthase; CBS; cardiant; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1..471
FT                   /*tag=  a
FT   CDS             472..1389
FT                   /*tag=  b
FT                   /product= "Human Prkag3 protein"
FT   3'UTR           1390..2109
FT                   /*tag=  c
XX
PN   WO200120003-A2.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-EP09896.
XX
PR   10-SEP-1999;   99EP-0402236.
PR   18-MAY-2000; 2000EP-0401388.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA   (ANDE/) ANDERSSON L.
PA   (LOOF/) LOOFT C.
PA   (KALM/) KALM E.
XX
PI   Andersson L,  Looft C,  Kalm E,  Milan D,  Robic A,  Rogel-Gaillard C;
PI   Iannuccelli N,  Gellin J,  Le Roy P,  Chardon P;
XX
DR   WPI; 2001-244810/25.
DR   P-PSDB; AAE00221.
XX
PT   New variants of the gamma subunit of vertebrate adenosine
PT   monophosphate-activated kinase for diagnosis or treatment of disorders
PT   associated with energy metabolism such as diabetes, obesity, and
PT   myopathy -
XX
PS   Claim 12; Fig 2; 71pp; English.
XX
CC   The present sequence is a cDNA encoding human adenosine monophosphate
CC   (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC   PRKAG3. Mutation in Prkag3 results in an altered regulation of
CC   carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is
CC   useful as therapeutic for treating carbohydrate metabolism disorders such
CC   as diabetes, obesity, and disorders associated with muscle metabolism
CC   such as myopathy and cardiovascular diseases, to modulate AMPK
CC   activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC   and its functionally altered mutants are useful for the diagnostic
CC   evaluation, genetic testing and prognosis of a metabolic disorder,
CC   preferably a carbohydrate metabolism disorder. Primers that can detect
CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
CC   from the expression of a functionally altered allele of PRKAG3.
CC   Transgenic animal and host cell transformed with PRKAG3 or a
CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC   screening compounds able to modulate AMPK activity. Nucleic acid
CC   encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
CC   of PRKAG3 and is useful in gene therapy.
XX
SQ   Sequence 2109 BP; 458 A; 621 C; 560 G; 470 T; 0 other;


  Query Match             98.4%;  Score 99.4;  DB 22;  Length 2109;
  Best Local Similarity   99.0%;  Pred. No. 4.1e-20;
  Matches  100;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      80 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 139

Qy      61 agggggaaggtccccggtccaggccagctgctgagtccacc 101
           |||||||||||||||||||||||| ||||||||||||||
Db     140 agggggaaggtccccggtccaggccaactgctgagtccacc 180


RESULT   3
AAD03320
ID   AAD03320 standard; cDNA; 2115 BP.
XX
```

```
AC   AAD03320;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Human AMPK gamma subunit muscle-specific isoform, complete PRKAG3 cDNA.
XX
KW   Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW   PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW   genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW   cystathione beta synthase; CBS; cardiant; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1395
FT                   /*tag= a
FT                   /product= "Human complete Prkag3 protein"
XX
PN   WO200120003-A2.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-EP09896.
XX
PR   10-SEP-1999;   99EP-0402236.
PR   18-MAY-2000; 2000EP-0401388.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA   (ANDE/) ANDERSSON L.
PA   (LOOF/) LOOFT C.
PA   (KALM/) KALM E.
XX
PI   Andersson L,  Looft C,  Kalm E,  Milan D,  Robic A,  Rogel-Gaillard C;
PI   Iannuccelli N,  Gellin J,  Le Roy P,  Chardon P;
XX
DR   WPI; 2001-244810/25.
DR   P-PSDB; AAE00223.
XX
PT   New variants of the gamma subunit of vertebrate adenosine
PT   monophosphate-activated kinase for diagnosis or treatment of disorders
PT   associated with energy metabolism such as diabetes, obesity, and
PT   myopathy  -
XX
PS   Claim 12; Page 65-68; 71pp; English.
XX
CC   The present sequence is a cDNA encoding human adenosine monophosphate
CC   (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC   complete PRKAG3. Mutation in Prkag3 results in an altered regulation of
CC   carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is
CC   useful as therapeutic for treating carbohydrate metabolism disorders such
CC   as diabetes, obesity, and disorders associated with muscle metabolism
CC   such as myopathy and cardiovascular diseases, to modulate AMPK
CC   activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC   and its functionally altered mutants are useful for the diagnostic
CC   evaluation, genetic testing and prognosis of a metabolic disorder,
CC   preferably a carbohydrate metabolism disorder. Primers that can detect
CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
CC   from the expression of a functionally altered allele of PRKAG3.
CC   Transgenic animal and host cell transformed with PRKAG3 or a
CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC   screening compounds able to modulate AMPK activity. Nucleic acid
CC   encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
CC   of PRKAG3 and is useful in gene therapy.
XX
SQ   Sequence 2115 BP; 460 A; 622 C; 562 G; 471 T; 0 other;

 Query Match             98.4%;  Score 99.4;  DB 22;  Length 2115;
 Best Local Similarity   99.0%;  Pred. No. 4.1e-20;
 Matches  100;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      86 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 145

Qy      61 agggggaaggtccccggtccaggccagctgctgagtccacc 101
           ||||||||||||||||||||||||| ||||||||||||||
Db     146 agggggaaggtccccggtccaggccaactgctgagtccacc 186


RESULT   4
AAD03295
ID   AAD03295 standard; cDNA; 1867 BP.
XX
AC   AAD03295;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Pig AMPK gamma subunit muscle-specific isoform, PRKAG3 cDNA.
XX
KW   Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW   PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW   genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW   cystathione beta synthase; CBS; cardiant; gene therapy; RN locus;
KW   chromosome 15; ss.
XX
OS   Sus scrofa.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1..471
FT                   /*tag= a
FT   CDS             472..1389
FT                   /*tag= b
FT                   /product= "Sus scrofa PRKAG3 protein"
FT   3'UTR           1390..1867
FT                   /*tag= c
XX
PN   WO200120003-A2.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-EP09896.
XX
PR   10-SEP-1999;   99EP-0402236.
PR   18-MAY-2000; 2000EP-0401388.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA   (ANDE/) ANDERSSON L.
PA   (LOOF/) LOOFT C.
PA   (KALM/) KALM E.
XX
PI   Andersson L,  Looft C,  Kalm E,  Milan D,  Robic A,  Rogel-Gaillard C;
PI   Iannuccelli N,  Gellin J,  Le Roy P,  Chardon P;
XX
DR   WPI; 2001-244810/25.
DR   P-PSDB; AAE00220.
XX
PT   New variants of the gamma subunit of vertebrate adenosine
PT   monophosphate-activated kinase for diagnosis or treatment of disorders
PT   associated with energy metabolism such as diabetes, obesity, and
PT   myopathy  -
XX
PS   Claim 12; Fig 2; 71pp; English.
XX
CC   The present sequence is a cDNA encoding pig adenosine monophosphate
CC   (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC   PRKAG3. Prkag3 gene is located in the RN locus of chromosome 15.
CC   Mutation in Prkag3 results in an altered regulation of carbohydrate
CC   metabolism, particularly in skeletal muscle. PRKAG3 is useful as
CC   therapeutic for treating carbohydrate metabolism disorders such as
CC   diabetes, obesity, and disorders associated with muscle metabolism
CC   such as myopathy and cardiovascular diseases, to modulate AMPK
CC   activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC   and its functionally altered mutants are useful for the diagnostic
```

```
CC   evaluation, genetic testing and prognosis of a metabolic disorder,
CC   preferably a carbohydrate metabolism disorder. Primers that can detect
CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
CC   from the expression of a functionally altered allele of PRKAG3.
CC   Transgenic animal and host cell transformed with PRKAG3 or a
CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC   screening compounds able to modulate AMPK activity. Nucleic acid
CC   encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
CC   of PRKAG3 and is useful in gene therapy.
XX
SQ   Sequence 1867 BP; 380 A; 583 C; 529 G; 375 T; 0 other;


  Query Match             71.5%;  Score 72.2;  DB 22;  Length 1867;
  Best Local Similarity   82.2%;  Pred. No. 4.1e-12;
  Matches   83;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 qggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           ||  ||| ||||  ||||||||||||||| ||   || ||||||||||| ||| || || |
Db      80 ggaacaaggcctctagatggacaaggcaggaggatgtagaggaaggggggcctccgggcc 139

Qy      61 agggggaaggtccccggtccaggccagctgctgagtccacc 101
             | |||||||||||| |||||||||| ||||||||||||
Db     140 cgagggaaggtccccagtccaggccagttgctgagtccacc 180


RESULT   5
AAD03319
ID   AAD03319 standard; cDNA; 1873 BP.
XX
AC   AAD03319;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Pig AMPK gamma subunit muscle-specific isoform, complete PRKAG3 cDNA.
XX
KW   Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW   PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW   genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW   cystathione beta synthase; CBS; cardiant; gene therapy; RN locus;
KW   chromosome 15; ss.
XX
OS   Sus scrofa.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1395
FT                   /*tag= a
FT                   /product= "Sus scrofa complete Prkag3 protein"
XX
PN   WO200120003-A2.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-EP09896.
XX
PR   10-SEP-1999;   99EP-0402236.
PR   18-MAY-2000; 2000EP-0401388.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA   (ANDE/) ANDERSSON L.
PA   (LOOF/) LOOFT C.
PA   (KALM/) KALM E.
XX
PI   Andersson L,  Looft C,  Kalm E,  Milan D,  Robic A,  Rogel-Gaillard C;
PI   Iannuccelli N,  Gellin J,  Le Roy P,  Chardon P;
XX
DR   WPI; 2001-244810/25.
DR   P-PSDB; AAE00222.
XX
PT   New variants of the gamma subunit of vertebrate adenosine
PT   monophosphate-activated kinase for diagnosis or treatment of disorders
PT   associated with energy metabolism such as diabetes, obesity, and
PT   myopathy -
XX
PS   Claim 12; Page 62-64; 71pp; English.
XX
CC   The present sequence is a cDNA encoding pig adenosine monophosphate
CC   (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC   complete PRKAG3. Prkag3 gene is located in the RN locus of chromosome
CC   15. Mutation in Prkag3 results in an altered regulation of carbohydrate
CC   metabolism, particularly in skeletal muscle. PRKAG3 is useful as
CC   therapeutic for treating carbohydrate metabolism disorders such as
CC   diabetes, obesity, and disorders associated with muscle metabolism
CC   such as myopathy and cardiovascular diseases, to modulate AMPK
CC   activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC   and its functionally altered mutants are useful for the diagnostic
CC   evaluation, genetic testing and prognosis of a metabolic disorder,
CC   preferably a carbohydrate metabolism disorder. Primers that can detect
CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
CC   from the expression of a functionally altered allele of PRKAG3.
CC   Transgenic animal and host cell transformed with PRKAG3 or a
CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC   screening compounds able to modulate AMPK activity. Nucleic acid
CC   encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
CC   of PRKAG3 and is useful in gene therapy.
XX
SQ   Sequence 1873 BP; 382 A; 580 C; 535 G; 376 T; 0 other;


  Query Match             71.5%;  Score 72.2;  DB 22;  Length 1873;
  Best Local Similarity   82.2%;  Pred. No. 4.1e-12;
  Matches   83;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           ||  ||| ||||  ||||||||||||||| ||   || ||||||||||| ||| || || |
Db      86 ggaacaaggcctctagatggacaaggcaggaggatgtagaggaaggggggcctccgggcc 145

Qy      61 agggggaaggtccccggtccaggccagctgctgagtccacc 101
             | |||||||||||| |||||||||| ||||||||||||
Db     146 cgagggaaggtccccagtccaggccagttgctgagtccacc 186


RESULT   6
AAD03321
ID   AAD03321 standard; DNA; 2022 BP.
XX
AC   AAD03321;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Sus scrofa PRKAG3 splice variant DNA.
XX
KW   Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW   PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW   genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW   cystathione beta synthase; CBS; cardiant; gene therapy; ds.
XX
OS   Sus scrofa.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1545
FT                   /*tag= a
FT                   /product= "Sus scrofa Prkag3 splice variant"
XX
PN   WO200120003-A2.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-EP09896.
XX
PR   10-SEP-1999;   99EP-0402236.
PR   18-MAY-2000; 2000EP-0401388.
```

```
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA   (ANDE/) ANDERSSON L.
PA   (LOOF/) LOOFT C.
PA   (KALM/) KALM E.
XX
PI   Andersson L,  Looft C,  Kalm E,  Milan D,  Robic A,  Rogel-Gaillard C;
PI   Iannuccelli N,  Gellin J,  Le Roy P,  Chardon P;
XX
DR   WPI; 2001-244810/25.
DR   P-PSDB; AAE00224.
XX
PT   New variants of the gamma subunit of vertebrate adenosine
PT   monophosphate-activated kinase for diagnosis or treatment of disorders
PT   associated with energy metabolism such as diabetes, obesity, and
PT   myopathy -
XX
PS   Claim 12; Page 69; 71pp; English.
XX
CC   The present sequence is pig adenosine monophosphate (AMP)-activated
CC   kinase (AMPK) gamma subunit muscle-specific isoform, PRKAG3 splice
CC   variant DNA. Prkag3 gene is located in the RN locus of chromosome 15.
CC   Mutation in Prkag3 results in an altered regulation of carbohydrate
CC   metabolism, particularly in skeletal muscle. PRKAG3 is useful as
CC   therapeutic for treating carbohydrate metabolism disorders such as
CC   diabetes, obesity, and disorders associated with muscle metabolism
CC   such as myopathy and cardiovascular diseases, to modulate AMPK
CC   activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC   and its functionally altered mutants are useful for the diagnostic
CC   evaluation, genetic testing and prognosis of a metabolic disorder,
CC   preferably a carbohydrate metabolism disorder. Primers that can detect
CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
CC   from the expression of a functionally altered allele of PRKAG3.
CC   Transgenic animal and host cell transformed with PRKAG3 or a
CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC   screening compounds able to modulate AMPK activity. Nucleic acid
CC   encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
CC   of PRKAG3 and is useful in gene therapy.
XX
SQ   Sequence 2022 BP; 412 A; 623 C; 593 G; 394 T; 0 other;

  Query Match             71.5%;  Score 72.2;  DB 22;  Length 2022;
  Best Local Similarity   82.2%;  Pred. No. 4.1e-12;
  Matches   83;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           ||  ||| ||||  ||||||||||||||| ||   || |||||||||| ||| || || |
Db     236 ggaacaaggcctctagatggacaaggcaggaggatgtagaggaaggggggcctccgggcc 295

Qy      61 agggggaaggtccccggtccaggccagctgctgagtccacc 101
             | |||||||||||| |||||||||| ||||||||||||
Db     296 cgagggaaggtccccagtccaggccagttgctgagtccacc 336


RESULT   7
AAH43682
ID   AAH43682 standard; DNA; 989 BP.
XX
AC   AAH43682;
XX
DT   21-JAN-2002  (first entry)
XX
DE   PRKAG3 intron 2 - intron 4.
XX
KW   Human; AMP-activated protein kinase gamma 3 subunit; PRKAG3; variant;
KW   metabolic disease; diabetes; obesity; substitution; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   intron          1..21
FT                   /*tag= a
FT                   /number = "Intron 2"
FT                   /note= "3' portion of intron 2"
FT   exon            22..177
FT                   /*tag= b
FT                   /number= "Exon 3"
FT   intron          178..541
FT                   /*tag= c
FT                   /number= "Intron 3"
FT   exon            542..945
FT                   /*tag= d
FT                   /number= "Exon 4"
FT   intron          946..989
FT                   /*tag= e
FT                   /number= "Intron 4"
FT                   /note= "5' portion of intron 4"
XX
PN   WO200177305-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-SE00765.
XX
PR   07-APR-2000; 2000US-195665P.
XX
PA   (AREX-) AREXIS AB.
XX
PI   Andersson L,  Luthman H,  Marklund S;
XX
DR   WPI; 2001-657170/75.
XX
PT   New variants of human AMP-activated protein kinase gamma3 subunit
PT   associated with a metabolic disease e.g. diabetes or obesity and method
PT   for determining a risk estimate of diseases in subject by detecting the
PT   variant -
XX
PS   Example 1; Fig 2; 25pp; English.
XX
CC   The sequences given in AAH43681-84 represents genomic fragments
CC   encoding the human AMP-activated protein kinase gamma 3 subunit
CC   (PRKAG3). Detecting the presence of the PRKAG3 DNA, or a variant,
CC   is useful in determining a risk estimate of a metabolic disease,
CC   such as diabetes or obesity, in a subject.  The variation may occur
CC   in exons 3, 4 or 10.  In exon 3 variation may be a substitution of
CC   a G for a C at nucleotide 320, resulting in the amino acid
CC   substitution P71A; in exon 4 variation may be a substitution of a
CC   T for a C at nucleotide 550; and in exon 10 variation may be a
CC   substitution of a T for a C at nucleotide 1037, resulting in the
CC   amino acid substitution R340W.  There may also be nucleotide variation
CC   in intron 6.
XX
SQ   Sequence 989 BP; 229 A; 306 C; 286 G; 168 T; 0 other;

  Query Match             70.3%;  Score 71;  DB 22;  Length 989;
  Best Local Similarity   100.0%;  Pred. No. 8.4e-12;
  Matches   71;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     109 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 168

Qy      61 agggggaaggt 71
           ||||||||||||
Db     169 agggggaaggt 179


RESULT   8
AAZ32059/c
ID   AAZ32059 standard; DNA; 8670 BP.
XX
AC   AAZ32059;
```

```
XX
DT   10-JAN-2000  (first entry)
XX
DE   Human METH2 related EST X16619.
XX
KW   Human; METH1; METH2; anti-angiogenic; metalloprotease thrombospondin;
KW   cancer; diagnosis; hyperproliferative disorder; autoimmune disease;
KW   angiogenesis inhibitor; abnormal wound healing; inflammation;
KW   rheumatoid arthritis; psoriasis; endometrial bleeding disorder;
KW   diabetic retinopathy; macula degeneration; haemangioma; detection;
KW   arterial-venous malformation; immune deficiency; ss.
XX
OS   Homo sapiens.
XX
PN   WO9937660-A1.
XX
PD   29-JUL-1999.
XX
PF   22-JAN-1999;   99WO-US01313.
XX
PR   23-JAN-1998;   98US-0072298.
PR   28-AUG-1998;   98US-0098539.
XX
PA   (IRUE/) IRUELA-ARISPE L.
PA   (HAST/) HASTINGS G A.
PA   (RUBE/) RUBEN S M.
XX
PI   Iruela-Arispe L,  Hastings GA,  Ruben SM;
XX
DR   WPI; 1999-590684/50.
XX
PT   New isolated metalloprotease thrombospondin polypeptides, useful for
PT   treating hyperproliferative disorders, cancers or autoimmune disorders
PT   -
XX
PS   Disclosure; Page 431-437; 457pp; English.
XX
CC   AAZ32000 and AAZ32001 encode, and AAY49501 and AAY49502 represent, human
CC   metalloprotease thrombospondin (METH) proteins METH1 and METH2
CC   respectively. METH1 and METH2 have been found to be potent inhibitors of
CC   angiogenesis both in vitro and in vivo. They can be used for treating
CC   cancer and other disorders related to angiogenesis including abnormal
CC   wound healing, inflammation, rheumatoid arthritis, psoriasis,
CC   endometrial bleeding disorders, diabetic retinopathy, some forms of
CC   macula degeneration, haemangiomas, and arterial-venous malformations.
CC   They may be useful in treating deficiencies or disorders of the immune
CC   system, by activating or inhibiting the proliferation, differentiation,
CC   or mobilisation (chemotaxis) of immune cells. The etiology of these
CC   immune deficiencies or disorders may be genetic, somatic, such as
CC   cancer or some autoimmune disorders, acquired (e.g. by chemotherapy or
CC   toxins), or infectious. They can also be used to treat inflammatory
CC   conditions, both chronic and acute conditions. The products can also be
CC   used for detection and diagnosis. AAZ32002 to AAZ32080, and AAY49503 to
CC   AAY49511 represent sequences given in the exemplification of the present
CC   invention.
XX
SQ   Sequence 8670 BP; 1624 A; 2421 C; 3266 G; 1359 T; 0 other;


  Query Match             31.5%;  Score 31.8;  DB 20;  Length 8670;
  Best Local Similarity   64.0%;  Pred. No. 3.8;
  Matches   48;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy      14 gagatggacaaggcagaagtcggtggaggaaggggagccaccaggtcaggggggaaggtcc 73
           ||| ||||   ||||| || ||||||||||| | |||| || | |       |  |  | ||
Db    6582 GAGGTGGAGGTGGCAGGAGGCGGTGGAGGAGGTGGAGGCAGCGGTATCATGAAACAGCCC 6523

Qy      74 ccggtccaggccagc 88
           || | |||  || ||
Db    6522 CCCGCCCAACCCCGC 6508


RESULT   9
```

```
AAC90316/c
ID   AAC90316 standard; DNA; 8670 BP.
XX
AC   AAC90316;
XX
DT   19-MAR-2001  (first entry)
XX
DE   X16619 cDNA clone.
XX
KW   METH; metalloprotease; thrombospondin; angiogenesis inhibition;
KW   cancer therapy; benign tumour; ocular angiogenic disease;
KW   rheumatoid arthritis; psoriasis; wound healing; endometriosis;
KW   vasculogenesis; granulation; hypertrophic scar; nonunion fracture;
KW   scleroderma, trachoma; vascular adhesion; myocardial angiogenesis;
KW   coronary collateral; cerebral collateral; arteriovenous malformation;
KW   ischaemic limb angiogenesis; Osler-Webber syndrome; wound granulation;
KW   plaque neovascularisation; telangiectasia; haemophiliac joint; EST;
KW   angiofibroma; fibromuscular dysplasia; expressed sequence tag;
KW   Crohn's disease; atherosclerosis; birth control; ss.
XX
OS   Unidentified.
XX
PN   WO200071577-A1.
XX
PD   30-NOV-2000.
XX
PF   25-MAY-2000; 2000WO-US14462.
XX
PR   25-MAY-1999;   99US-0318208.
PR   20-JUL-1999;   99US-0144882.
PR   10-AUG-1999;   99US-0147823.
PR   13-AUG-1999;   99US-0373658.
PR   22-DEC-1999;   99US-0171503.
PR   22-FEB-2000; 2000US-0183792.
XX
PA   (HUMA-) HUMAN GENOME SCI INC.
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.
PA   (IRUE/) IRUELA-ARISPE L.
PA   (HAST/) HASTINGS G A.
PA   (RUBE/) RUBEN S M.
PA   (JONA/) JONAK Z L.
PA   (TRUL/) TRULLI S H.
PA   (FORN/) FORNWALD J A.
PA   (TERR/) TERRETT J A.
XX
PI   Iruela-Arispe L,  Hastings GA,  Ruben SM,  Jonak ZL,  Trulli SH;
PI   Fornwald JA,  Terrett JA;
XX
DR   WPI; 2001-025136/03.
XX
PT   METH1 and METH2 polynucleotides and encoded polypeptides, used to
PT   inhibit angiogensis in the treatment of disorders such as cancer,
PT   rheumatoid arthritis and psoriasis -
XX
PS   Claim 14; Pages 731-737; 768pp; English.
XX
CC   The present invention relates to human METH1 and METH2, (ME for
CC   metalloprotease and TH for thrombospondin; see AAB50002 and AAB50003).
CC   The present sequence is an expressed sequence tag (EST) for METH. METH
CC   can be used for inhibiting angiogenesis in an individual, and for
CC   treating cancer, benign tumours, an ocular angiogenic disease,
CC   rheumatoid arthritis, psoriasis, delayed wound healing, endometriosis,
CC   vasculogenesis, granulations, hypertrophic scars, nonunion fractures,
CC   scleroderma, trachoma, vascular adhesions, myocardial angiogenesis,
CC   coronary collaterals, cerebral collaterals, arteriovenous malformations,
CC   ischaemic limb angiogenesis, Osler-Webber syndrome, plaque
CC   neovascularisation, telangiectasia, haemophiliac joints, angiofibroma,
CC   fibromuscular dysplasia, wound granulation, Crohn's disease or
CC   atherosclerosis. METH can also be used in birth control. METH can also
CC   be used in diagnostic methods for the prognosis of cancer.
XX
SQ   Sequence 8670 BP; 1624 A; 2421 C; 3266 G; 1359 T; 0 other;
```

```
  Query Match             31.5%;  Score 31.8;  DB 22;  Length 8670;
  Best Local Similarity   64.0%;  Pred. No. 3.8;
  Matches   48;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy      14 gagatggacaaggcagaagtcggtggaggaaggggagccaccaggtcaggggggaaggtcc 73
           ||| ||||   ||||| || |||||||||| | |||| || | |       |  |  | ||
Db    6582 GAGGTGGAGGTGGCAGGAGGCGGTGGAGGAGGTGGAGGCAGCGGTATCATGAAACAGCCC 6523

Qy      74 ccggtccaggccagc 88
           || | |||  || ||
Db    6522 CCCGCCCAACCCCGC 6508
```

```
RESULT  10
AAS84420
ID   AAS84420 standard; cDNA; 711 BP.
XX
AC   AAS84420;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #20224.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG20233.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 20224; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 711 BP; 139 A; 227 C; 249 G; 96 T; 0 other;
```

```
  Query Match             29.7%;  Score 30;  DB 23;  Length 711;
  Best Local Similarity   67.7%;  Pred. No. 9.2;
  Matches   42;  Conservative    0;  Mismatches   20;  Indels    0;  Gaps    0;

Qy      30 aagtcggtggaggaaggggagccaccaggtcaggggggaaggtccccggtccaggccagct 89
           || | | |||||||||| |||  | || || ||  |||||  ||| |||   ||  | |||
Db     431 aactggctggaggaagaggacaccccggggcaaagggaacctccacggagaagaactgct 490

Qy      90 gc 91
           ||
Db     491 gc 492
```

```
RESULT  11
AAS84523
ID   AAS84523 standard; cDNA; 2541 BP.
XX
AC   AAS84523;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #20327.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG20336.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 20327; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
```

```
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2541 BP; 476 A; 780 C; 787 G; 498 T; 0 other;


  Query Match             29.5%;  Score 29.8;  DB 23;  Length 2541;
  Best Local Similarity   63.0%;  Pred. No. 13;
  Matches   46;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy      15 agatggacaaggcagaagtcggtggaggaaggggagccaccaggtcaggggaaggtccc 74
           ||| |  |  ||||| | |||||| | ||  |  |||  ||||||||||    ||||
Db    1104 agaggcccgtggcaggggccggtggggaaactgtggcccggaggtcaggggtcctgtcca 1163

Qy      75 cggtccaggccag 87
           | |||  ||| |
Db    1164 cagtcaccgcctg 1176


RESULT  12
AAD22004
ID   AAD22004 standard; cDNA; 4125 BP.
XX
AC   AAD22004;
XX
DT   12-FEB-2002  (first entry)
XX
DE   Human transporters and ion channels (TRICH)-12 cDNA.
XX
KW   Human; transporter and ion channel; TRICH; akinesia; cystic fibrosis;
KW   diabetes mellitus; Parkinson's disease; myasthenia gravis; dementia;
KW   cardiac disorder; angina; hypertension; myocarditis; hyperglycaemia;
KW   neurological disorder; Alzheimer's disease; cataract; infertility;
KW   Wilson's disease; schizophrenia; Grave's disease; addison's disease;
KW   Huntington's disease; multiple sclerosis; meningitis; hypotensive;
KW   cardiant; nootropic; neuroprotective; neuroleptic; ophthalmological;
KW   antithyroid; anticonvulsant; goitre; antiinflammatory; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3933
FT                   /*tag= a
FT                   /product= "Human transporters and ion channels
FT                   (TRICH)-12"
XX
PN   WO200177174-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-US11206.
XX
PR   06-APR-2000; 2000US-195595P.
PR   12-APR-2000; 2000US-196872P.
PR   20-APR-2000; 2000US-199020P.
PR   28-APR-2000; 2000US-200552P.
PR   05-MAY-2000; 2000US-202348P.
PR   11-MAY-2000; 2000US-203495P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Reddy R,  Thornton M,  Borowsky ML,  Tang YT,  Khan FA,  Tribouley CM;
PI   Gandhi AR,  Yao MG,  Sanjanwala MS,  Baughn MR,  Nguyen DB;
PI   Policky JL,  Yue H,  Seilhamer JJ,  Walia NK,  Lal P,  Kearney L;
PI   Walsh RT,  Lu DAM,  Lu Y,  Greene BD,  Raumann BE,  Patterson C;
XX
DR   WPI; 2002-017448/02.
DR   P-PSDB; AAE13285.
XX
PT   Polypeptides of human transporters and ion channels, useful for
PT   diagnosing, treating or preventing disorders of transport,
PT   neurological, muscle, immunological and cell proliferative disorders -
XX
PS   Claim 5; Page 147-148; 150pp; English.
XX
CC   The invention relates to human transporters and ion channels (TRICH)
CC   and the polynucleotides encoding them. The composition comprising TRICH
CC   or agonist of TRICH is useful for treating a disease or condition
CC   associated with decreased expression of functional TRICH or condition
CC   associated with overexpression of TRICH respectively. The composition
CC   comprising Ab is useful for diagnosing a condition of disease associated
CC   with expression of TRICH in a subject, where the disorders include a
CC   transport disorder such as akinesia, cystic fibrosis, diabetes mellitus,
CC   Parkinson's disease, myasthenia gravis, cardiac disorders associated
CC   with transport e.g. angina, hypertension, myocarditis, neurological
CC   disorders associated with transport e.g. Alzheimer's disease, Wilson's
CC   disease, schizophrenia, cataracts, infertility, hyperglycaemia, Grave's
CC   disease, goitre, addison's disease, Huntington's disease, dementia,
CC   multiple sclerosis, bacterial and viral meningitis. TRICH DNA is useful
CC   for generating a transcript image of a tissue or cell type, which
CC   represents the global pattern of gene expression by a particular tissue
CC   or cell type and for analysing the proteome of a tissue or cell type.
CC   TRICH DNA is used in gene therapy. The present sequence is human
CC   TRICH12 cDNA.
XX
SQ   Sequence 4125 BP; 989 A; 1109 C; 1188 G; 839 T; 0 other;


  Query Match             29.5%;  Score 29.8;  DB 24;  Length 4125;
  Best Local Similarity   60.5%;  Pred. No. 13;
  Matches   49;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy       7 aagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtcaggggg 66
           |||| | | |||||| ||       |  | ||  ||  |||||| | |  | ||  ||||
Db    3901 aagcgtcgtgatggaaaaatgtcctggagctgatggctggggaggcccaggatcccgggg 3960

Qy      67 aaggtccccggtccaggccag 87
           |   |||  || ||||| |||
Db    3961 atcctcctgggaccaggacag 3981


RESULT  13
AAH06841/c
ID   AAH06841 standard; cDNA; 886 BP.
XX
AC   AAH06841;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA clone (5'-primer) SEQ ID NO:3676.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
```

```
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 1; SEQ ID 3676; 2537pp + CD ROM; English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 886 BP; 169 A; 259 C; 213 G; 241 T; 4 other;


  Query Match             29.1%;  Score 29.4;  DB 22;  Length 886;
  Best Local Similarity   58.6%;  Pred. No. 14;
  Matches   51;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy       2 ggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtca 61
           ||| || | ||    | || | |||| |    ||| ||||     ||||  | |||| |
Db     100 GGCTAAGGACTCCGTAGGGGCCAGGCGGGCAACGGCGGAGCGCTCGGAGGGAGCAGGGCC 41

Qy      62 gggggaaggtccccggtccaggccagc 88
             | |  | ||||||||  |  ||| ||
Db      40 AGAGAGAAGTCCCCGGGGCCAGCCTGC 14


RESULT  14
AAH14551/c
ID   AAH14551 standard; cDNA; 1457 BP.
XX
AC   AAH14551;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:12117.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000;  2000JP-0118776.
PR   02-MAY-2000;  2000JP-0183767.
PR   09-JUN-2000;  2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 8; SEQ ID 12117; 2537pp + CD ROM; English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 1457 BP; 291 A; 442 C; 365 G; 359 T; 0 other;


  Query Match             29.1%;  Score 29.4;  DB 22;  Length 1457;
  Best Local Similarity   58.6%;  Pred. No. 15;
  Matches   51;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy       2 ggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtca 61
           ||| || | ||    | || | |||| |    ||| ||||     ||||  | |||| |
Db     100 GGCTAAGGACTCCGTAGGGGCCAGGCGGGCAACGGCGGAGCGCTCGGAGGGAGCAGGGCC 41

Qy      62 gggggaaggtccccggtccaggccagc 88
             | |  | ||||||||  |  ||| ||
Db      40 AGAGAGAAGTCCCCGGGGCCAGCCTGC 14


RESULT  15
AAV59752/c
ID   AAV59752 standard; DNA; 1482 BP.
XX
AC   AAV59752;
XX
DT   19-JAN-1999  (first entry)
XX
DE   Human secreted protein gene 96 clone HAQBK61.
XX
KW   Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW   diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW   developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW   immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW   inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW   cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW   osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW   endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS   Homo sapiens.
XX
```

PN WO9839448-A2.
XX
PD 11-SEP-1998.
XX
PF 06-MAR-1998; 98WO-US04493.
XX
PR 02-OCT-1997; 97US-0061060.
PR 07-MAR-1997; 97US-0038621.
PR 07-MAR-1997; 97US-0040161.
PR 07-MAR-1997; 97US-0040162.
PR 07-MAR-1997; 97US-0040163.
PR 07-MAR-1997; 97US-0040333.
PR 07-MAR-1997; 97US-0040334.
PR 07-MAR-1997; 97US-0040336.
PR 07-MAR-1997; 97US-0040626.
PR 11-APR-1997; 97US-0043311.
PR 11-APR-1997; 97US-0043312.
PR 11-APR-1997; 97US-0043313.
PR 11-APR-1997; 97US-0043314.
PR 11-APR-1997; 97US-0043568.
PR 11-APR-1997; 97US-0043569.
PR 11-APR-1997; 97US-0043576.
PR 11-APR-1997; 97US-0043578.
PR 11-APR-1997; 97US-0043580.
PR 11-APR-1997; 97US-0043669.
PR 11-APR-1997; 97US-0043670.
PR 11-APR-1997; 97US-0043671.
PR 11-APR-1997; 97US-0043672.
PR 11-APR-1997; 97US-0043674.
PR 23-MAY-1997; 97US-0047492.
PR 23-MAY-1997; 97US-0047500.
PR 23-MAY-1997; 97US-0047501.
PR 23-MAY-1997; 97US-0047502.
PR 23-MAY-1997; 97US-0047503.
PR 23-MAY-1997; 97US-0047581.
PR 23-MAY-1997; 97US-0047582.
PR 23-MAY-1997; 97US-0047583.
PR 23-MAY-1997; 97US-0047584.
PR 23-MAY-1997; 97US-0047585.
PR 23-MAY-1997; 97US-0047586.
PR 23-MAY-1997; 97US-0047587.
PR 23-MAY-1997; 97US-0047588.
PR 23-MAY-1997; 97US-0047589.
PR 23-MAY-1997; 97US-0047590.
PR 23-MAY-1997; 97US-0047592.
PR 23-MAY-1997; 97US-0047593.
PR 23-MAY-1997; 97US-0047594.
PR 23-MAY-1997; 97US-0047595.
PR 23-MAY-1997; 97US-0047596.
PR 23-MAY-1997; 97US-0047597.
PR 23-MAY-1997; 97US-0047598.
PR 23-MAY-1997; 97US-0047599.
PR 23-MAY-1997; 97US-0047600.
PR 23-MAY-1997; 97US-0047601.
PR 23-MAY-1997; 97US-0047612.
PR 23-MAY-1997; 97US-0047613.
PR 23-MAY-1997; 97US-0047614.
PR 23-MAY-1997; 97US-0047615.
PR 23-MAY-1997; 97US-0047617.
PR 23-MAY-1997; 97US-0047618.
PR 23-MAY-1997; 97US-0047632.
PR 23-MAY-1997; 97US-0047633.
PR 06-JUN-1997; 97US-0048964.
PR 06-JUN-1997; 97US-0048974.
PR 13-JUN-1997; 97US-0049610.
PR 08-JUL-1997; 97US-0051926.
PR 16-JUL-1997; 97US-0052874.
PR 18-AUG-1997; 97US-0055724.
PR 22-AUG-1997; 97US-0056630.
PR 22-AUG-1997; 97US-0056631.
PR 22-AUG-1997; 97US-0056632.
PR 22-AUG-1997; 97US-0056636.
PR 22-AUG-1997; 97US-0056637.
PR 22-AUG-1997; 97US-0056662.
PR 22-AUG-1997; 97US-0056664.
PR 22-AUG-1997; 97US-0056845.
PR 22-AUG-1997; 97US-0056862.
PR 22-AUG-1997; 97US-0056864.
PR 22-AUG-1997; 97US-0056872.
PR 22-AUG-1997; 97US-0056874.
PR 22-AUG-1997; 97US-0056875.
PR 22-AUG-1997; 97US-0056876.
PR 22-AUG-1997; 97US-0056877.
PR 22-AUG-1997; 97US-0056878.
PR 22-AUG-1997; 97US-0056879.
PR 22-AUG-1997; 97US-0056880.
PR 22-AUG-1997; 97US-0056881.
PR 22-AUG-1997; 97US-0056882.
PR 22-AUG-1997; 97US-0056884.
PR 22-AUG-1997; 97US-0056886.
PR 22-AUG-1997; 97US-0056887.
PR 22-AUG-1997; 97US-0056888.
PR 22-AUG-1997; 97US-0056889.
PR 22-AUG-1997; 97US-0056892.
PR 22-AUG-1997; 97US-0056893.
PR 22-AUG-1997; 97US-0056894.
PR 22-AUG-1997; 97US-0056903.
PR 22-AUG-1997; 97US-0056908.
PR 22-AUG-1997; 97US-0056909.
PR 22-AUG-1997; 97US-0056910.
PR 22-AUG-1997; 97US-0056911.
PR 05-SEP-1997; 97US-0057650.
PR 05-SEP-1997; 97US-0057669.
PR 05-SEP-1997; 97US-0057761.
PR 12-SEP-1997; 97US-0058785.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA;
PI Feng P, Ferrie AM, Fischer CL, Florence KA, Greene JM, Hu JS;
PI Kyaw H, Lafleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA;
PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
XX
DR WPI; 1998-506364/43.
DR P-PSDB; AAW74967.
XX
PT New isolated human genes and the secreted polypeptide(s) they encode
PT - useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders
XX
PS Claim 1; Page 478-479; 721pp; English.
XX
CC This sequence represents a nucleic acid molecule designated Gene 96 from
CC the human cDNA clone HAQBK61 (deposited as clone ATCC 97897 and ATCC
CC 209043) which encodes a secreted human protein. The gene can be used to
CC generate fusion proteins by linking to the gene to a human
CC immunoglobulin Fc portion (e.g. AAV59502) for increasing the stability of
CC the fused protein as compared to the human protein only.
CC The invention relates to 186 novel genes and their fragments (nucleic
CC acid sequences: AAV59511-V59812; amino acid sequences AAW74731-W75026)
CC which are useful for preventing, treating or ameliorating medical
CC conditions e.g. by protein or gene therapy. Also, pathological
CC conditions can be diagnosed by determining the amount of the new
CC polypeptides in a sample or by determining the presence of mutations in
CC the new polynucleotides. Specific uses are described for each of the 186
CC polynucleotides, based on which tissues they are most highly expressed in
CC (see AAV59511 for described uses).
XX
SQ Sequence 1482 BP; 308 A; 444 C; 368 G; 362 T; 0 other;

Query Match 29.1%; Score 29.4; DB 19; Length 1482;
Best Local Similarity 58.6%; Pred. No. 15;
Matches 51; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

Qy 2 ggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtca 61

```
          ||| || | ||    | || | |||| |    ||| ||||     ||||  | |||| |
Db      90 GGCTAAGGACTCCGTAGGGGCCAGGCGGGCAACGGCGGAGCGCTCGGAGGGAGCAGGGCC 31

Qy      62 qqqqqaaggtccccggtccaggccagc 88
             | |  | ||||||||  |  ||| ||
Db      30 AGAGAGAAGTCCCCGGGGCCAGCCTGC 4
```

Search completed: October 3, 2002, 16:30:43
Job time: 14313 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 16:15:01 ; Search time 5701.1 Seconds
(without alignments)
239.110 Million cell updates/sec

Title: US-09-826-581-5_COPY_180_280
Perfect score: 101
Sequence: 1 gggccaaagccttgagatgg.....ggccagctgctgagtccacc 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	67.4	66.7	399	9	AW356079	AW356079 38073 MAR
2	67.4	66.7	422	9	AW427435	AW427435 63185 MAR
3	67.4	66.7	621	9	BB628877	BB628877 BB628877
4	67.4	66.7	655	9	BB629521	BB629521 BB629521
5	65.8	65.1	685	9	BB630381	BB630381 BB630381
6	54.6	54.1	548	10	BI775360	BI775360 467815 MA
7	53.8	53.3	516	9	AI664508	AI664508 uk25b05.y
8	52.6	52.1	444	10	BF890374	BF890374 291826 MA
c 9	35.6	35.2	922	10	BI754053	BI754053 603027607
c 10	32.6	32.3	406	9	BB759831	BB759831 BB759831
c 11	32.6	32.3	457	9	BB780864	BB780864 BB780864
c 12	32.6	32.3	736	10	BI765906	BI765906 603046160
13	32.4	32.1	151	10	BG244994	BG244994 602358260
14	32	31.7	454	9	BB860266	BB860266 BB860266
15	32	31.7	479	10	BF549656	BF549656 UI-R-C2-n
c 16	31.2	30.9	896	10	BF526844	BF526844 602070405
c 17	31	30.7	327	9	BB753946	BB753946 BB753946
18	31	30.7	1074	10	BG966577	BG966577 602834407
c 19	30.8	30.5	365	9	AA787203	AA787203 m8d07a1.r
20	30.8	30.5	491	10	BF651383	BF651383 274155 MA
c 21	30.8	30.5	722	12	AZ188883	AZ188883 SP_1013_A
22	30.6	30.3	512	9	AI939950	AI939950 CM0-CT001
23	30.6	30.3	1195	10	BE299786	BE299786 600944223
24	30.4	30.1	370	10	BG086356	BG086356 H3124H09-
25	30.4	30.1	443	9	BB845787	BB845787 BB845787
c 26	30.4	30.1	514	10	BI181375	BI181375 UNL-P-FN-
27	30.2	29.9	167	10	BF393187	BF393187 UI-R-CA0-
c 28	30.2	29.9	211	9	BB438733	BB438733 BB438733
29	30.2	29.9	428	10	BF394635	BF394635 UI-R-CA0-
c 30	30.2	29.9	545	9	AW413109	AW413109 uq52e07.x
c 31	30.2	29.9	825	10	BI910794	BI910794 603070386
32	30.2	29.9	1079	12	CNS05JOV	AL340456 Tetraodon
c 33	30	29.7	250	10	BE426030	BE426030 WHE0327_E
c 34	30	29.7	409	10	BE426831	BE426831 WHE0333_H
c 35	30	29.7	429	10	BE489634	BE489634 WHE1071-1
c 36	30	29.7	1253	11	AK015101	AK015101 Mus muscu
c 37	29.8	29.5	335	10	Z25209	Z25209 HSBA5G012 S
38	29.8	29.5	385	12	AQ852870	AQ852870 LMAJFV1_1
c 39	29.8	29.5	581	12	BH414079	BH414079 1007036C0
c 40	29.8	29.5	662	12	CNS048SD	AL279670 Tetraodon
41	29.8	29.5	740	10	BF153550	BF153550 028D07 Ma
42	29.8	29.5	1101	12	CNS000O4	AL066116 Drosophil
c 43	29.6	29.3	285	9	BB039663	BB039663 BB039663
44	29.6	29.3	375	9	AW298840	AW298840 332 MARC
c 45	29.6	29.3	431	9	BB822055	BB822055 BB822055

ALIGNMENTS

RESULT 1
AW356079
LOCUS AW356079 399 bp mRNA linear EST 25-APR-2001
DEFINITION 38073 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION AW356079
VERSION AW356079.1 GI:6860085
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 399)
AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keele,J.W.
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
JOURNAL Genome Res. 11 (4), 626-630 (2001)
MEDLINE 21180013
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 20 and -minmatch 12 options.
PCR PRimers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCCAGTCACGACG
Plate: 17 row: P column: 9
Seq primer: ATTTAGGTGACACTATAG.
FEATURES Location/Qualifiers
source 1. .399
/organism="Bos taurus"
/db_xref="taxon:9913"

```
                     /clone_lib="MARC 2BOV"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled tissue from testis, thymus,
                     semitendonosus muscle, longissimus muscle, pancreas,
                     adrenal, and endometrium."
BASE COUNT       100 a    126 c    117 g     56 t
ORIGIN


  Query Match             66.7%;  Score 67.4;  DB 9;  Length 399;
  Best Local Similarity   79.2%;  Pred. No. 1.8e-07;
  Matches   80;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           || |||| | ||  ||||||| ||| ||| ||   || |||||| |||||| |||||
Db     122 GGACCAAGGTCTCAAGATGGAAAAGCCAGGAGGATGTAGAGGAAAGGGAGCTGCCAGGCC 181

Qy      61 agggggaaggtccccggtccaggccagctgctgagtccacc 101
             || || |||||||| ||||||| ||||||||||||||||
Db     182 TGGAGGGAGGTCCCCAGTCCAGGGCAGCTGCTGAGTCCACC 222


RESULT   2
AW427435
LOCUS       AW427435                422 bp    mRNA    linear   EST 25-APR-2001
DEFINITION  63185 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   AW427435
VERSION     AW427435.1  GI:6955382
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 422)
  AUTHORS   Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
            ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
            Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
            Keele,J.W.
  TITLE     Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
  JOURNAL   Genome Res. 11 (4), 626-630 (2001)
  MEDLINE   21180013
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 20
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 32  row: F  column: 22
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .422
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="MARC 3BOV"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled tissue from marrow, alveolar
                     macrophage, ovary, fetal semitendonosus muscle, and fetal
                     longissimus muscle."
BASE COUNT       102 a    136 c    125 g     59 t
ORIGIN
```

```
  Query Match             66.7%;  Score 67.4;  DB 9;  Length 422;
  Best Local Similarity   79.2%;  Pred. No. 1.8e-07;
  Matches   80;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           || |||| | ||  ||||||| ||| ||| ||   || |||||| |||||| |||||
Db     170 GGACCAAGGTCTCAAGATGGAAAAGCCAGGAGGATGTAGAGGAAAGGGAGCTGCCAGGCC 229

Qy      61 agggggaaggtccccggtccaggccagctgctgagtccacc 101
             || || |||||||| ||||||| ||||||||||||||||
Db     230 TGGAGGGAGGTCCCCAGTCCAGGGCAGCTGCTGAGTCCACC 270


RESULT   3
BB628877
LOCUS       BB628877                621 bp    mRNA    linear   EST 31-AUG-2001
DEFINITION  BB628877 RIKEN full-length enriched, 16 days neonate cerebellum Mus
            musculus cDNA clone 9630039L22 5', mRNA sequence.
ACCESSION   BB628877
VERSION     BB628877.1  GI:15399483
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 621)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A., Aizawa
            ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K., Arakawa,T.,
            Ishii,Y. and Hayashizaki,Y.
             Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct.
            Func. Genomics 2 pre, L72-L86 (2001
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .621
```

/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="9630039L22"
/clone_lib="RIKEN full-length enriched, 16 days neonate cerebellum"
/tissue_type="cerebellum"
/dev_stage="16 days neonate"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 20.0 and subtraction to Rot = 370.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."
BASE COUNT 143 a 194 c 183 g 101 t
ORIGIN

Query Match 66.7%; Score 67.4; DB 9; Length 621;
Best Local Similarity 79.2%; Pred. No. 2e-07;
Matches 80; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
Db 203 GAGTCAAGGCTTCCAGATGGACGAGACAGGAGGCCGTAGAGGAAGCAGAACCACCAGGTT 262

Qy 61 agggggaaggtccccggtccaggccagctgctgagtccacc 101
Db 263 TGGGAGAAGGTGCCCAGTCCAGACCAGCTGCTGAGTCCACC 303

RESULT 4
BB629521
LOCUS BB629521 655 bp mRNA linear EST 31-AUG-2001
DEFINITION BB629521 RIKEN full-length enriched, adult male bone Mus musculus cDNA clone 9830138C07 5', mRNA sequence.
ACCESSION BB629521
VERSION BB629521.1 GI:15399646
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 655)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. . 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. . 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K., Arakawa,T., Ishii,Y. and Hayashizaki,Y.
Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct. Func. Genomics 2 pre, L72-L86 (2001
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
e mouse tissues.
FEATURES Location/Qualifiers
source 1. .655
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="9830138C07"
/clone_lib="RIKEN full-length enriched, adult male bone"
/sex="male"
/tissue_type="bone"
/dev_stage="adult"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."
BASE COUNT 148 a 204 c 202 g 101 t
ORIGIN

Query Match 66.7%; Score 67.4; DB 9; Length 655;
Best Local Similarity 79.2%; Pred. No. 2e-07;
Matches 80; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
Db 264 GAGTCAAGGCTTCCAGATGGACGAGACAGGAGGCCGTAGAGGAAGCAGAACCACCAGGTT 323

Qy 61 agggggaaggtccccggtccaggccagctgctgagtccacc 101
Db 324 TGGGAGAAGGTGCCCAGTCCAGACCAGCTGCTGAGTCCACC 364

RESULT 5
BB630381
LOCUS BB630381 685 bp mRNA linear EST 26-OCT-2001
DEFINITION BB630381 RIKEN full-length enriched, 6 days neonate skin Mus

musculus cDNA clone A030014A04 5', mRNA sequence.
ACCESSION BB630381
VERSION BB630381.1 GI:16467364
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 685)
AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.
TITLE RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
JOURNAL Unpublished (2001)
COMMENT Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. . 10 (10), 1617-1630 (2000)
wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. . 10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and Hayashizaki,Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
e mouse tissues.
FEATURES Location/Qualifiers
source 1. .685
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="A030014A04"
/clone_lib="RIKEN full-length enriched, 6 days neonate skin"
/tissue_type="skin"
/dev_stage="6 days neonate"
/lab_host="DH10B"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 185.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. "
BASE COUNT 165 a 199 c 198 g 123 t
ORIGIN

Query Match 65.1%; Score 65.8; DB 9; Length 685;
Best Local Similarity 78.2%; Pred. No. 5.1e-07;
Matches 79; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

Qy 1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
Db 170 GAGTCAAGGCTTCCAGATGGACGAGACAGGAGGCCGTAGAGGAAGCAGAACCACCAGGTT 229

Qy 61 agggggaaggtccccggtccaggccagctgctgagtccacc 101
Db 230 TGGGAGAAGGTGCCCAGTCCATACCAGCTGCTGAGTCCACC 270

RESULT 6
BI775360
LOCUS BI775360 548 bp mRNA linear EST 25-SEP-2001
DEFINITION 467815 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BI775360
VERSION BI775360.1 GI:15776346
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 548)
AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keele,J.W.
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
JOURNAL Genome Res. 11 (4), 626-630 (2001)
MEDLINE 21180013
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR PRimers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCCAGTCACGACG
Plate: 90 row: I column: 16
Seq primer: ATTTAGGTGACACTATAG.
FEATURES Location/Qualifiers
source 1. .548
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone_lib="MARC 2BOV"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI; Library made from pooled tissue from testis, thymus, semitendonosus muscle, longissimus muscle, pancreas, adrenal, and endometrium."
BASE COUNT 119 a 173 c 163 g 93 t
ORIGIN

```
  Query Match           54.1%;  Score 54.6;  DB 10;  Length 548;
  Best Local Similarity  77.6%;  Pred. No. 0.00033;
  Matches   66;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      17 atggacaaggcagaagtcggtggaggaaggggagccaccaggtcaggggggaaggtccccg 76
           ||||| ||| ||| ||    ||  ||||| ||||||  ||||| | || || ||||||||
Db      33 ATGGAAAAGCCAGGAGGATGTAAAGGAAAGGGAGCTGCCAGGCCTGGAGGGAGGTCCCCA 92

Qy      77 gtccaggccagctgctgagtccacc 101
           |||| | ||||||||||||||||||
Db      93 TTCCAAGGCAGCTGCTGAGTCCACC 117


RESULT   7
AI664508
LOCUS       AI664508                516 bp    mRNA    linear   EST 10-MAY-1999
DEFINITION  uk25b05.y1 Sugano mouse embryo mewa Mus musculus cDNA clone
            IMAGE:1970001 5', mRNA sequence.
ACCESSION   AI664508
VERSION     AI664508.1  GI:4768091
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 516)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Other_ESTs: uk25b05.x1
            Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:986741
            Seq primer: custom primer used
            High quality sequence stop: 484.
FEATURES             Location/Qualifiers
     source          1. .516
                     /organism="Mus musculus"
                     /strain="C57BL"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:1970001"
                     /clone_lib="Sugano mouse embryo mewa"
                     /dev_stage="embryo, 14 dpc"
                     /lab_host="DH10B"
                     /note="Vector: pME18S-FL3; Site_1: DraIII (CACTGTGTG);
                     Site_2: DraIII (CACCATGTG); 1st strand cDNA was primed
                     with an oligo(dT) primer [ATGTGGCCTTTTTTTTTTTTTTTT];
                     double-stranded cDNA was ligated to a DraIII adaptor
                     [TGTTGGCCTACTGG], digested and cloned into distinct DraIII
                     sites of the pME18S-FL3 vector (5' site CACTGTGTG, 3' site
                     CACCATGTG). XhoI should be used to isolate the cDNA
                     insert. Size selection was performed to exclude fragments
                     <1.5kb. Library constructed by Dr. Sumio Sugano
                     (University of Tokyo Institute of Medical Science).
                     Custom primers for sequencing: 5' end primer
                     CTTCTGCTCTAAAAGCTGCG and 3' end primer
                     CGACCTGCAGCTCGAGCACA."
BASE COUNT      113 a    139 c    171 g     90 t      3 others
ORIGIN


  Query Match           53.3%;  Score 53.8;  DB 9;  Length 516;
  Best Local Similarity  77.2%;  Pred. No. 0.00051;
  Matches   78;  Conservative    0;  Mismatches   22;  Indels    1;  Gaps    1;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           | | ||| || |  |||||||| || ||| || | || |||||||  || |||||||||
Db     408 GAGTCAAGGCTTCCAGATGGACGAGACAGGAGGCCGTAGAGGAAGCAGAACCACCAGGTT 467

Qy      61 agggggaaggtccccggtccaggccagctgctgagtccacc 101
           ||| || ||| ||  |||||| ||||||||||||||||||
Db     468 TGGGAGAGGGTGCCA-GTCCAGACCAGCTGCTGAGTCCACC 507


RESULT   8
BF890374
LOCUS       BF890374                444 bp    mRNA    linear   EST 25-APR-2001
DEFINITION  291826 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BF890374
VERSION     BF890374.1  GI:12281760
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 444)
  AUTHORS   Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
            ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
            Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
            Keele,J.W.
  TITLE     Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
  JOURNAL   Genome Res. 11 (4), 626-630 (2001)
  MEDLINE   21180013
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 57  row: H  column: 11
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .444
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="MARC 3BOV"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled tissue from marrow, alveolar
                     macrophage, ovary, fetal semitendonosus muscle, and fetal
                     longissimus muscle."
BASE COUNT       91 a    147 c    139 g     67 t
ORIGIN


  Query Match           52.1%;  Score 52.6;  DB 10;  Length 444;
  Best Local Similarity  81.3%;  Pred. No. 0.00099;
  Matches   61;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy      27 cagaagtcggtggaggaaggggagccaccaggtcaggggggaaggtccccggtccaggcca 86
           ||| ||   || |||||| ||||||  ||||| | || || |||||||| ||||||| ||
Db       1 CAGGAGGATGTAGAGGAAAGGGAGCTGCCAGGCCTGGAGGGAGGTCCCCAGTCCAGGGCA 60

Qy      87 gctgctgagtccacc 101
           |||||||||||||||
```

```
Db      61 GCTGCTGAGTCCACC 75


RESULT   9
BI754053/c
LOCUS       BI754053                922 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  603027607F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5198016 5',
            mRNA sequence.
ACCESSION   BI754053
VERSION     BI754053.1  GI:15745631
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 922)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11495  row: n  column: 01
            High quality sequence stop: 830.
FEATURES             Location/Qualifiers
     source          1. .922
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5198016"
                     /clone_lib="NIH_MGC_114"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
                     male brains, age range 23-27 yo. Library is oligo-dT
                     primed and directionally cloned (EcoRV site is  destroyed
                     upon cloning).  Average insert size 1.5 kb, insert size
                     range 1-3 kb.  Library is normalized and enriched for
                     full-length  clones and was constructed by C. Gruber
                     (Invitrogen).  Research  Genetics tracking code 019. Note:
                     this is a NIH_MGC Library."
BASE COUNT      216 a    268 c    284 g    154 t
ORIGIN


  Query Match             35.2%;  Score 35.6;  DB 10;  Length 922;
  Best Local Similarity   62.2%;  Pred. No. 23;
  Matches   56;  Conservative    0;  Mismatches   34;  Indels    0;  Gaps    0;

Qy       4 ccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtcagg 63
           | | ||  |||    |||   |   ||| ||   || || |  ||||||| |||| |
Db     915 CGACAGGTTTGCCCCGGATGGGTTCGAATTCCCAGGTGGGACTGGAGCCAACAGGGCCCT 856

Qy      64 gggaaggtccccggtccaggccagctgctg 93
           ||||  |  |||| ||| ||||| ||||||
Db     855 GGGAGTGGGCCCGTTCCTGGCCACCTGCTG 826


RESULT  10
BB759831/c
LOCUS       BB759831                406 bp    mRNA    linear   EST 17-OCT-2001
DEFINITION  BB759831 RIKEN full-length enriched, melanocyte Mus musculus cDNA
            clone G270121L23 3', mRNA sequence.
ACCESSION   BB759831
VERSION     BB759831.1  GI:16206159
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 406)
  AUTHORS   Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
            Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
            ,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
            Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.
            , Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
            Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
            ,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.
            , Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
            2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .406
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="G270121L23"
                     /clone_lib="RIKEN full-length enriched, melanocyte"
                     /cell_type="melanocyte"
BASE COUNT       77 a    130 c     98 g    101 t
ORIGIN


  Query Match             32.3%;  Score 32.6;  DB 9;  Length 406;
  Best Local Similarity   58.9%;  Pred. No. 1.1e+02;
  Matches   56;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy       7 aagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtcaggggg 66
           ||||| |||| |   ||||    | |  || ||  ||| || ||   ||||    || ||
Db     257 AAGCCCTGAGTTCCCCAAGCAGAAGGCAGGAGGGAGAAAGGCAGAGGCCAGAGTGGGAGG 198

Qy      67 aaggtccccggtccaggccagctgctgagtccacc 101
             |||  ||| |  |||| | |||    ||| || |
Db     197 CAGGGACCCCGAGCAGGTCTGCTATCAAGTTCAGC 163


RESULT  11
BB780864/c
LOCUS       BB780864                457 bp    mRNA    linear   EST 15-NOV-2001
DEFINITION  BB780864 RIKEN full-length enriched, colon RCB-0549 Cle-H3 cDNA Mus
            musculus cDNA clone G430068O10 3', mRNA sequence.
```

```
ACCESSION   BB780864
VERSION     BB780864.1  GI:16941564
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 457)
  AUTHORS   Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
            Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
            ,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
            Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.
            , Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
            Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
            ,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.
            , Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
            2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .457
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="G430068O10"
                     /clone_lib="RIKEN full-length enriched, colon RCB-0549
                     Cle-H3 cDNA"
                     /tissue_type="colon"
                     /cell_line="RCB-0549 Cle-H3"
BASE COUNT       93 a    145 c    106 g    113 t
ORIGIN
```

```
  Query Match             32.3%;  Score 32.6;  DB 9;  Length 457;
  Best Local Similarity   58.9%;  Pred. No. 1.1e+02;
  Matches   56;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy       7 aagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtcaggggg 66
           ||||| |||| |   ||||    | |  || ||  ||| || ||   |||     || ||
Db     313 AAGCCCTGAGTTCCCCAAGCAGAAGGCAGGAGGGAGAAAGGCAGAGGCCACAGTGGGAGG 254

Qy      67 aaggtccccggtccaggccagctgctgagtccacc 101
           ||||  ||| |  |||| | |||    ||| || |
Db     253 AAGGGACCCCGAGCAGGTCTGCTATCAAGTTCAGC 219
```

```
RESULT  12
BI765906/c
LOCUS       BI765906                736 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  603046160F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5186367 5',
            mRNA sequence.
ACCESSION   BI765906
VERSION     BI765906.1  GI:15757484
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 736)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM11465  row: h  column: 16
            High quality sequence stop: 736.
FEATURES             Location/Qualifiers
     source          1. .736
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5186367"
                     /clone_lib="NIH_MGC_116"
                     /lab_host="DH10B"
                     /note="Organ: pooled colon, kidney, stomach; Vector:
                     pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
                     source anonymous pool of 3 colons, age 26 yo male, 49 yo
                     female, 71 yo male colon; 46 yo male kidney, and pool of 2
                     stomachs, 62 yo male and 70 yo female.  Library is
                     oligo-dT primed and  directionally cloned (EcoRV site is
                     destroyed upon cloning).  Average insert size 1.4 kb,
                     insert size range 1-3 kb.  Library is normalized and
                     enriched for full-length clones and was constructed by C.
                     Gruber  (Invitrogen).  Research Genetics tracking code
                     023. Note: this is a NIH_MGC Library."
BASE COUNT      138 a    213 c    194 g    191 t
ORIGIN
```

```
  Query Match             32.3%;  Score 32.6;  DB 10;  Length 736;
  Best Local Similarity   60.9%;  Pred. No. 1.2e+02;
  Matches   53;  Conservative    0;  Mismatches   34;  Indels    0;  Gaps    0;

Qy       2 ggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtca 61
             ||| || | ||    | || | |||| |    ||| ||||     ||||| |||||| |
Db     113 GGCTAAGGACTCCGTAGGGGCCAGGCGGGCAACGGCGGAGCGCTCGGAGCGACCAGGGCC 54

Qy      62 gggggaaggtccccggtccaggccagc 88
             | |  | ||||||||  |  ||| ||
Db      53 AGAGAGAAGTCCCCGGGGCCAGCCTGC 27
```

```
RESULT  13
BG244994
LOCUS       BG244994                151 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  602358260F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4486719 5',
            mRNA sequence.
ACCESSION   BG244994
VERSION     BG244994.1  GI:12754809
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 151)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10330  row: h  column: 16
            High quality sequence stop: 151.
FEATURES             Location/Qualifiers
     source          1. .151
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4486719"
                     /clone_lib="NCI_CGAP_Mam1"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="10 months, virgin"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
BASE COUNT       21 a      38 c      70 g      22 t
ORIGIN


  Query Match             32.1%;  Score 32.4;  DB 10;  Length 151;
  Best Local Similarity   68.2%;  Pred. No. 97;
  Matches   45;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy      14 gagatggacaaggcagaagtcggtggaggaaggggagccaccaggtcagggggaaggtcc 73
           ||| ||||| |||||| | |||||||||| | ||| | || ||   ||||| | | |
Db      66 GAGGTGGACTCGGCAGACGCCGGTGGAGGGTGAGGATCACCCGGGAAGGGGGGGACGACG 125

Qy      74 ccggtc 79
           ||||
Db     126 TGGGTC 131


RESULT  14
BB860266
LOCUS       BB860266                 454 bp    mRNA    linear   EST 26-NOV-2001
DEFINITION  BB860266 RIKEN full-length enriched, pooled cell lines Mus musculus
            cDNA clone G430012004 5', mRNA sequence.
ACCESSION   BB860266
VERSION     BB860266.1  GI:17101720
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 454)
  AUTHORS   Akimura,T., Arakawa,T., Carninci,P., Furuno,M., Hanagaki,T.,
            Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T., Imotani,K., Ishii
            ,Y., Ito,M., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Matsuyama,T.,
            Nakamura,M., Nishi,K., Nomura,K., Numasaki,R., Okazaki,Y., Okido,T.
            , Saito,R., Sakai,C., Sakai,K., Sakazume,N., Sasaki,D., Sato,K.,
            Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagawa
            ,A., Takahashi,F., Takaku-Akahira,S., Tanaka,T., Tomaru,A., Toya,T.
            , Watahiki,A., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura,T., et al.
            2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .454
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="G430012004"
                     /clone_lib="RIKEN full-length enriched, pooled cell lines"
                     /note="pooled cell lines ; (cell_line=CRL-1751 WEHI 164),
                     (cell_line=CRL-2116 JC), (cell_line=RCB-0035 WEHI-3),
                     (cell_line=RCB-0464 Meth-A), (cell_line=RCB-0545 OHTA),
                     (cell_line=RCB-0559 K-1 . F1), (cell_line=RCB-1283 B16
                     melanoma), (cell_type=B cells, cell_line=CRL-1702 WEHI 231
                     ), (cell_type=Leydig cells, cell_line=CRL-2065 MLTC-1),
                     (cell_type=Nullipotent stem cell, cell_line=CRL-2070 NE),
                     (tissue_type=bladder, cell_line=RCB-0544 MBT-2),
                     (tissue_type=bone marrow, cell_type=stroma cell,
                     cell_line= CRL-2028 SR-4987), (tissue_type=colon,
                     cell_line=RCB-0549 Cle-H3), (tissue_type=kidney,
                     cell_line=CCL-142 RAG), (tissue_type=submandibular gland,
                     cell_line=CRL-1734 SCA-9 clone 15), (strain=BALB/C,
                     cell_type=B cells, cell_line=CRL-1669 BCL1 Clone 13.20-3B3
                     ), (strain=C3H, tissue_type=brain, cell_line=CRL-1443
                     BC3H1)"
BASE COUNT      121 a     123 c     133 g      77 t
ORIGIN


  Query Match             31.7%;  Score 32;  DB 9;  Length 454;
  Best Local Similarity   60.2%;  Pred. No. 1.6e+02;
  Matches   53;  Conservative    0;  Mismatches   35;  Indels    0;  Gaps    0;

Qy       6 aaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtcagggg 65
           | | | | ||||  | |||||||| |  || |   | | |||||| || ||  | |  |
Db     357 ACACCATAGAGAATGTCAAGGCAAAGATCAGGACAAGGAGGGGATCCCCCCTGACCAGCA 416

Qy      66 gaaggtccccggtccaggccagctgctg 93
           || | |   |  | ||||| ||| ||||
Db     417 GAGGCTGATCTTTGCAGGCAAGCAGCTG 444


RESULT  15
BF549656
LOCUS       BF549656                 479 bp    mRNA    linear   EST 12-DEC-2000
DEFINITION  UI-R-C2-ng-f-01-0-UI.r1 UI-R-C2 Rattus norvegicus cDNA clone
            UI-R-C2-ng-f-01-0-UI 5', mRNA sequence.
ACCESSION   BF549656
```

```
VERSION     BF549656.1  GI:11659386
KEYWORDS    EST.
SOURCE      Norway rat.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 479)
  AUTHORS   Bonaldo,M.F., Lennon,G. and Soares,M.B.
  TITLE     Normalization and subtraction: two approaches to facilitate gene
            discovery
  JOURNAL   Genome Res. 6 (9), 791-806 (1996)
  MEDLINE   97044477
COMMENT     Contact: Soares, MB
            Program for Rat Gene Discovery and Mapping
            University of Iowa
            451 Eckstein Medical Research Building Iowa City, IA 52242, USA
            Tel: 319 335 8250
            Fax: 319 335 9565
            Email: msoares@blue.weeg.uiowa.edu
            cDNA Library Preparation: M.B. Soares Lab Clone distribution:
            clones will be available through Research Genetics (www.resgen.com)
            This clone is also available through the I.M.A.G.E. Consortium at
            LLNL (info@image.llnl.gov). IMAGE ID= 1776352
            Seq primer: M13 Forward.
FEATURES             Location/Qualifiers
     source          1. .479
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /clone="UI-R-C2-ng-f-01-0-UI"
                     /clone_lib="UI-R-C2"
                     /dev_stage="adult"
                     /lab_host="DH10B (Life Technologies)"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not 1; Site_2: Eco RI; The UI-R-C2
                     library is a subtracted library derived from the UI-R-C1
                     library, which is a subtracted library derived from the
                     UI-R-C0 library.  The UI-R-C0 library consisted of a
                     mixture of individually tagged normalized  libraries
                     constructed from rat placenta, adult lung, brain, liver,
                     kidney,  heart, spleen, ovary, muscle, 8, 12 and 18-day
                     embryo. The tag is a string of  3-5 nucleotides present
                     between the Not I site and the oligo-dT track which allows
                     identification of the library of origin of a clone within
                     the mixture. The subtracted library (UI-R-C2) was
                     constructed as follows: PCR amplified cDNA inserts from
                     UI-R-C1 clones from which 3' ESTs had been derived was
                     used as a driver in a hybridization with the UI-R-C1
                     library in the form of single-stranded circles. The
                     remaining single-stranded circles (subtracted library) was
                     purified by hydroxyapatite column chromatography,
                     converted to double-stranded circles and electroporated
                     into DH10B bacteria (Life Technologies) to generate the
                     UI-R-C2 library. This procedure has been previously
                     described (Bonaldo, Lennon and Soares, Genome Research 6:
                     791-806, 1996)"
BASE COUNT      103 a    163 c    146 g     65 t      2 others
ORIGIN
```

```
  Query Match             31.7%;  Score 32;  DB 10;  Length 479;
  Best Local Similarity   62.5%;  Pred. No. 1.6e+02;
  Matches   50;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy       9 gccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtcagggggaa 68
           | |||| | || |||||||| ||   ||| ||   | |||    ||||||  || ||
Db      48 GACTTGGGCTGCACAAGGCACCAGAGCGTGCAGACTGAGGATGGCCCAGGTGGGGTGGCC 107

Qy      69 ggtccccggtccaggccagc 88
             |  ||| | ||| | ||||
Db     108 AGAGCCCTGGCCAAGGCAGC 127
```

Search completed: October 3, 2002, 16:15:04
Job time: 16959 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 16:22:08 ; Search time 180.77 Seconds
(without alignments)
137.241 Million cell updates/sec

Title: US-09-826-581-5_COPY_180_280
Perfect score: 101
Sequence: 1 gggccaaagccttgagatgg.....ggccagctgctgagtccacc 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	28.8	28.5	6503	4	US-09-404-650-12	Sequence 12, Appl
	2	28.6	28.3	268	3	US-08-800-840-1	Sequence 1, Appli
	3	28.6	28.3	268	3	US-08-776-859-1	Sequence 1, Appli
	4	28.6	28.3	268	5	PCT-US95-09905-1	Sequence 1, Appli
	5	28.6	28.3	428	3	US-08-826-532-1	Sequence 1, Appli
	6	28.6	28.3	428	4	US-09-228-324A-1	Sequence 1, Appli
	7	28.6	28.3	596	2	US-08-105-989-3	Sequence 3, Appli
	8	28.6	28.3	596	3	US-09-138-922-3	Sequence 3, Appli
	9	28.6	28.3	2500	1	US-08-550-715-10	Sequence 10, Appl
	10	28.6	28.3	6060	5	PCT-US96-09430-7	Sequence 7, Appli
c	11	26.8	26.5	1140	1	US-08-289-653-2	Sequence 2, Appli
c	12	26.8	26.5	2845	1	US-08-289-653-1	Sequence 1, Appli
c	13	26.8	26.5	5100	1	US-08-164-292B-15	Sequence 15, Appl
c	14	26.8	26.5	5100	1	US-08-164-292B-17	Sequence 17, Appl
c	15	26.8	26.5	5100	1	US-08-164-292B-19	Sequence 19, Appl
c	16	26.8	26.5	5100	1	US-08-164-292B-21	Sequence 21, Appl
c	17	26.8	26.5	5100	1	US-08-164-292B-23	Sequence 23, Appl
c	18	26.8	26.5	5100	1	US-08-164-292B-25	Sequence 25, Appl
c	19	26.8	26.5	5100	3	US-08-845-623-15	Sequence 15, Appl
c	20	26.8	26.5	5100	3	US-08-845-623-17	Sequence 17, Appl
c	21	26.8	26.5	5100	3	US-08-845-623-19	Sequence 19, Appl
c	22	26.8	26.5	5100	3	US-08-845-623-21	Sequence 21, Appl
c	23	26.8	26.5	5100	3	US-08-845-623-23	Sequence 23, Appl
c	24	26.8	26.5	5100	3	US-08-845-623-25	Sequence 25, Appl
c	25	26.8	26.5	5100	3	US-08-815-927-15	Sequence 15, Appl
c	26	26.8	26.5	5100	3	US-08-815-927-17	Sequence 17, Appl
c	27	26.8	26.5	5100	3	US-08-815-927-19	Sequence 19, Appl
c	28	26.8	26.5	5100	3	US-08-815-927-21	Sequence 21, Appl
c	29	26.8	26.5	5100	3	US-08-815-927-23	Sequence 23, Appl
c	30	26.8	26.5	5100	3	US-08-815-927-25	Sequence 25, Appl
c	31	26.8	26.5	5100	4	US-09-103-330-15	Sequence 15, Appl
c	32	26.8	26.5	5100	4	US-09-103-330-17	Sequence 17, Appl
c	33	26.8	26.5	5100	4	US-09-103-330-19	Sequence 19, Appl
c	34	26.8	26.5	5100	4	US-09-103-330-21	Sequence 21, Appl
c	35	26.8	26.5	5100	4	US-09-103-330-23	Sequence 23, Appl
c	36	26.8	26.5	5100	4	US-09-103-330-25	Sequence 25, Appl
c	37	26.8	26.5	34446	4	US-09-103-330-35	Sequence 35, Appl
	38	26.2	25.9	458	4	US-09-328-111-464	Sequence 464, App
c	39	25.6	25.3	4403765	4	US-09-103-840A-2	Sequence 2, Appli
c	40	25.6	25.3	4411529	4	US-09-103-840A-1	Sequence 1, Appli
c	41	25.4	25.1	4035	6	5198359-1	Patent No. 5198359
c	42	25.4	25.1	4035	6	5449756-1	Patent No. 5449756
c	43	25.2	25.0	833	2	US-08-343-443B-97	Sequence 97, Appl
	44	25	24.8	215	1	US-08-207-481-44	Sequence 44, Appl
	45	25	24.8	215	5	PCT-US95-02689-51	Sequence 51, Appl

ALIGNMENTS

```
RESULT   1
US-09-404-650-12/c
; Sequence 12, Application US/09404650
; Patent No. 6309858
; GENERAL INFORMATION:
;  APPLICANT: Dietrich, Paul S.
;  APPLICANT: McGivern, Joseph G.
;  TITLE OF INVENTION: T-TYPE CALCIUM CHANNEL VARIANTS; COMPOSITIONS THEREOF;
;  TITLE OF INVENTION: AND USES
;  FILE REFERENCE: R0043B-REG sequence listing
;  CURRENT APPLICATION NUMBER: US/09/404,650
;  CURRENT FILING DATE: 1999-09-23
;  NUMBER OF SEQ ID NOS: 12
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
;   LENGTH: 6503
;   TYPE: DNA
;   ORGANISM: Rattus sp.
US-09-404-650-12


  Query Match            28.5%;  Score 28.8;  DB 4;  Length 6503;
  Best Local Similarity  60.0%;  Pred. No. 4.1;
  Matches   48;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy      22 caaggcagaagtcggtggaggaaggggagccaccaggtcaggggggaaggtccccggtcca 81
           ||  |  | ||  |  || | | ||| ||  |||||| ||||||  ||| |   || | |
Db    5474 CACAGGGGCAGGGGCAGGGGCAGGGGCAGGGACCAGGGCAGGGGCCAGGGCAGGGGCCGA 5415

Qy      82 ggccagctgctgagtccacc 101
           |||||   ||    |||| |
Db    5414 GGCCATGGGCCATCTCCAGC 5395


RESULT   2
US-08-800-840-1
; Sequence 1, Application US/08800840
; Patent No. 6060288
;  GENERAL INFORMATION:
;    APPLICANT:  Adams, Christopher P.
;    APPLICANT:  Boles, Truett C.
;    APPLICANT:  Muir, Andrew R.
;    APPLICANT:  Kron, Stephen J.
;    TITLE OF INVENTION:  Method and Apparatus for Performing
;    TITLE OF INVENTION:  Amplification of Nucleic Acis on Supports
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Hamilton, Brook, Smith & Reynolds, P.C.
;      STREET:  Two Militia Drive
;      CITY:  Lexington
```

```
;       STATE:  Massachusetts
;       COUNTRY:  US
;       ZIP:  02173
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/800,840
;       FILING DATE:  14-FEB-1997
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/285,385
;       FILING DATE:  03-AUG-1994
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/09905
;       FILING DATE:  03-AUG-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/776,859
;       FILING DATE:  03-FEB-1997
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Brook, David E.
;       REGISTRATION NUMBER:  22,592
;       REFERENCE/DOCKET NUMBER:  MST97-01A2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617-861-6240
;       TELEFAX:  617-861-9540
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  268 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
US-08-800-840-1


  Query Match           28.3%;  Score 28.6;  DB 3;  Length 268;
  Best Local Similarity   55.6%;  Pred. No. 2.6;
  Matches   55;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           | |||| | ||| | | ||| |||    |       || | ||| |||| ||   ||||| |
Db      50 GAGCCACACCCTAGGGTTGGCCAATCTACTCCCAGGAGCAGGGAGGGCAGGAGCCAGGGC 109

Qy      61 agggggaaggtccccggtccaggccagctgctgagtcca 99
             |||    |      | || |    |||| ||  ||   | ||
Db     110 TGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACA 148


RESULT   3
US-08-776-859-1
; Sequence 1, Application US/08776859
; Patent No. 6090592
; GENERAL INFORMATION:
;  APPLICANT: Mosaic Technologies, Inc.
;  APPLICANT: Adams, Christopher P.
;  APPLICANT: Kron, Stephen J.
;  TITLE OF INVENTION: Method and Apparatus for Performing
;  TITLE OF INVENTION: Amplification of Nucleic Acids on Supports
;  FILE REFERENCE: Mosaic
;  CURRENT APPLICATION NUMBER: US/08/776,859
;  CURRENT FILING DATE: 1997-05-29
;  EARLIER APPLICATION NUMBER: 08/285,385
;  EARLIER FILING DATE: 1994-03-08
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 268
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-08-776-859-1


  Query Match           28.3%;  Score 28.6;  DB 3;  Length 268;
  Best Local Similarity   55.6%;  Pred. No. 2.6;
  Matches   55;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           | |||| | ||| | | ||| |||    |       || | ||| |||| ||   ||||| |
Db      50 GAGCCACACCCTAGGGTTGGCCAATCTACTCCCAGGAGCAGGGAGGGCAGGAGCCAGGGC 109

Qy      61 agggggaaggtccccggtccaggccagctgctgagtcca 99
             |||    |      | || |    |||| ||  ||   | ||
Db     110 tgggcataaaagtcagggcagagccatctattgcttaca 148


RESULT   4
PCT-US95-09905-1
; Sequence 1, Application PC/TUS9509905
;  GENERAL INFORMATION:
;     APPLICANT:
;     APPLICANT:
;     APPLICANT:
;     APPLICANT:
;     TITLE OF INVENTION:  METHOD AND APPARATUS FOR PERFORMING
;     TITLE OF INVENTION:  AMPLIFICATION OF NUCLEIC ACIDS ON SUPPORTS
;     NUMBER OF SEQUENCES:  4
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  WOLF, GREENFIELD & SACKS, P.C.
;       STREET:  600 ATLANTIC AVENUE
;       CITY:  BOSTON
;       STATE:  MASSACHUSETTS
;       COUNTRY:  USA
;       ZIP:  02210
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/09905
;       FILING DATE:  FILED HEREWITH
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/285,385
;       FILING DATE:  03-AUGUST-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  JANIUK, ANTHONY J.
;       REGISTRATION NUMBER:  29,809
;       REFERENCE/DOCKET NUMBER:  B0833/7001WO
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  617-720-3500
;       TELEFAX:  617-720-2441
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  268 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     ORIGINAL SOURCE:
;       ORGANISM:  Homo sapiens
PCT-US95-09905-1


  Query Match           28.3%;  Score 28.6;  DB 5;  Length 268;
  Best Local Similarity   55.6%;  Pred. No. 2.6;
  Matches   55;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           | |||| | ||| | | ||| |||    |       || | ||| |||| ||   ||||| |
Db      50 GAGCCACACCCTAGGGTTGGCCAATCTACTCCCAGGAGCAGGGAGGGCAGGAGCCAGGGC 109

Qy      61 agggggaaggtccccggtccaggccagctgctgagtcca 99
```

```
          |||   |     | || |   |||| ||  ||  | ||
Db    110 TGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACA 148

RESULT   5
US-08-826-532-1
; Sequence 1, Application US/08826532B
; Patent No. 6027923
; GENERAL INFORMATION:
;  APPLICANT: Wallace, Robert B.
;  TITLE OF INVENTION: Linked Linear Amplification of Nucleic Acids
;  FILE REFERENCE: 3239-102P
;  CURRENT APPLICATION NUMBER: US/08/826,532B
;  CURRENT FILING DATE: 1997-04-02
;  EARLIER APPLICATION NUMBER: US 08/475,605
;  EARLIER FILING DATE: 1995-06-07
;  NUMBER OF SEQ ID NOS: 20
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 428
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-08-826-532-1


  Query Match           28.3%;  Score 28.6;  DB 3;  Length 428;
  Best Local Similarity  55.6%;  Pred. No. 2.8;
  Matches   55;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           | |||| | ||| | | ||| |||   |      || | ||| |||| ||   ||||| |
Db      13 gagccacaccctagggttggccaatctactcccaggagcagggagggcaggagccagggc 72

Qy      61 agggggaaggtccccggtccaggccagctgctgagtcca 99
             |||   |     | || |   |||| ||  ||  | ||
Db      73 tgggcataaaagtcagggcagagccatctattgcttaca 111


RESULT   6
US-09-228-324A-1
; Sequence 1, Application US/09228324A
; Patent No. 6335184
; GENERAL INFORMATION:
;  APPLICANT: Reyes, Antonio A
;  APPLICANT:  Wallace, Robert B.
;  APPLICANT:  Ugozzoli, Luis A.
;  TITLE OF INVENTION: Linked Linear Amplification of Nucleic Acids
;  FILE REFERENCE: 3239-103P
;  CURRENT APPLICATION NUMBER: US/09/228,324A
;  CURRENT FILING DATE:  1999-01-11
;  PRIOR APPLICATION NUMBER: US 08/826,532
;  PRIOR FILING DATE: 1997-04-02
;  NUMBER OF SEQ ID NOS: 64
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 428
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-228-324A-1


  Query Match           28.3%;  Score 28.6;  DB 4;  Length 428;
  Best Local Similarity  55.6%;  Pred. No. 2.8;
  Matches   55;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           | |||| | ||| | | ||| |||   |      || | ||| |||| ||   ||||| |
Db      13 gagccacaccctagggttggccaatctactcccaggagcagggagggcaggagccagggc 72

Qy      61 agggggaaggtccccggtccaggccagctgctgagtcca 99
             |||   |     | || |   |||| ||  ||  | ||
Db      73 tgggcataaaagtcagggcagagccatctattgcttaca 111


RESULT   7
US-08-105-989-3
; Sequence 3, Application US/08105989
; Patent No. 5922854
;  GENERAL INFORMATION:
;    APPLICANT:  Kumar, Ramesh
;    APPLICANT:  Sharma, Ajay
;    APPLICANT:  Paulhiac, Clara
;    APPLICANT:  Khoury-Christianson, Anastasia P.
;    APPLICANT:  Midha, Sunita
;    TITLE OF INVENTION:  Production of Human Hemoglobin in
;    TITLE OF INVENTION:  Transgenic Pigs.
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PENNIE & EDMONDS
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/105,989
;      FILING DATE:  11-AUG-1993
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  6794-030
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-8864/9741
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  596 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  DNA (genomic)
US-08-105-989-3


  Query Match           28.3%;  Score 28.6;  DB 2;  Length 596;
  Best Local Similarity  55.6%;  Pred. No. 3;
  Matches   55;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           | |||| | ||| | | ||| |||   |      || | ||| |||| ||   ||||| |
Db     431 GAGCCACACCCTAGGGTTGGCCAATCTACTCCCAGGAGCAGGGAGGGCAGGAGCCAGGGC 490

Qy      61 agggggaaggtccccggtccaggccagctgctgagtcca 99
             |||   |     | || |   |||| ||  ||  | ||
Db     491 TGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACA 529


RESULT   8
US-09-138-922-3
; Sequence 3, Application US/09138922
; Patent No. 6147202
;  GENERAL INFORMATION:
;    APPLICANT:  Kumar, Ramesh
;    APPLICANT:  Sharma, Ajay
;    APPLICANT:  Paulhiac, Clara
;    APPLICANT:  Khoury-Christianson, Anastasia P.
;    APPLICANT:  Midha, Sunita
;    TITLE OF INVENTION:  Production of Human Hemoglobin in
;    TITLE OF INVENTION:  Transgenic Pigs.
```

```
;    NUMBER OF SEQUENCES:  26
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PENNIE & EDMONDS
;      STREET:  1155 Avenue of the Americas
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  U.S.A.
;      ZIP:  10036-2711
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Releasc #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/138,922
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/105,989
;      FILING DATE:  11-AUG-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Coruzzi, Laura A.
;      REGISTRATION NUMBER:  30,742
;      REFERENCE/DOCKET NUMBER:  6794-030
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 790-9090
;      TELEFAX:  (212) 869-8864/9741
;      TELEX:  66141 PENNIE
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  596 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  unknown
;    MOLECULE TYPE:  DNA (genomic)
US-09-138-922-3


  Query Match             28.3%;  Score 28.6;  DB 3;  Length 596;
  Best Local Similarity   55.6%;  Pred. No. 3;
  Matches   55;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           | |||| | ||| | | ||| |||   |       || | ||| |||| ||   ||||| |
Db     431 GAGCCACACCCTAGGGTTGGCCAATCTACTCCCAGGAGCAGGGAGGGCAGGAGCCAGGGC 490

Qy      61 aggggaaggtccccggtccaggccagctgctgagtcca 99
           |||   |     | || |   |||| ||  ||  | ||
Db     491 TGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACA 529


RESULT   9
US-08-550-715-10
; Sequence 10, Application US/08550715
; Patent No. 5750345
;  GENERAL INFORMATION:
;    APPLICANT:  Bowie, Lemuel J.
;    TITLE OF INVENTION:  Human '-Thalassemia Mutations as a Predictor of
;    TITLE OF INVENTION:  Blood-Related Disorders
;    NUMBER OF SEQUENCES:  13
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/550,715
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Gass, David A.
;      REGISTRATION NUMBER:  38,153
;      REFERENCE/DOCKET NUMBER:  28493/32834
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312/474-6300
;      TELEFAX:  312/474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2500 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  join(687..778, 909..1131, 1982..2107)
US-08-550-715-10


  Query Match             28.3%;  Score 28.6;  DB 1;  Length 2500;
  Best Local Similarity   55.6%;  Pred. No. 3.9;
  Matches   55;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           | |||| | ||| | | ||| |||   |       || | ||| |||| ||   ||||| |
Db     541 GAGCCACACCCTAGGGTTGGCCAATCTACTCCCAGGAGCAGGGAGGGCAGGAGCCAGGGC 600

Qy      61 aggggaaggtccccggtccaggccagctgctgagtcca 99
           |||   |     | || |   |||| ||  ||  | ||
Db     601 TGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACA 639


RESULT  10
PCT-US96-09430-7
; Sequence 7, Application PC/TUS9609430
;  GENERAL INFORMATION:
;    APPLICANT:  Glazer, Peter M.
;    TITLE OF INVENTION:  TREATMENT OF HEMOGLOBINOPATHIES
;    NUMBER OF SEQUENCES:  23
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OncorPharm, Inc.
;      STREET:  200 Perry Parkway
;      CITY:  Gaithersburg
;      STATE:  Maryland
;      COUNTRY:  US
;      ZIP:  20877
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US96/09430
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/473,845
;      FILING DATE:  07-JUN-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Karta, Glenn E.
;      REGISTRATION NUMBER:  30,649
;      REFERENCE/DOCKET NUMBER:  PA-0040
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301-527-2058
;      TELEFAX:  301-208-6997
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6060 base pairs
```

```
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     ORIGINAL SOURCE:
;       ORGANISM:  Homo sapiens
;     POSITION IN GENOME:
;       CHROMOSOME/SEGMENT:  11
PCT-US96-09430-7


  Query Match             28.3%;  Score 28.6;  DB 5;  Length 6060;
  Best Local Similarity   55.6%;  Pred. No. 4.7;
  Matches   55;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy       1 gggccaaagccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtc 60
           | |||| | ||| | | ||| |||   |      || | ||| |||| ||   ||||| |
Db    2041 GAGCCACACCCTAGGGTTGGCCAATCTACTCCCAGGAGCAGGGAGGGCAGGAGCCAGGGC 2100

Qy      61 agggggaaggtccccggtccaggccagctgctgagtcca 99
             |||   |     | || |   |||| ||  ||  | ||
Db    2101 TGGGCATAAAAGTCAGGGCAGAGCCATCTATTGCTTACA 2139


RESULT  11
US-08-289-653-2/c
; Sequence 2, Application US/08289653
; Patent No. 5543322
;  GENERAL INFORMATION:
;    APPLICANT:  Kazuaki KITANO et al.
;    TITLE OF INVENTION:  DNA AND ITS USE
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wenderoth, Lind & Ponack
;      STREET:  805 Fifteenth Street, N.W., #700
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 500 kb
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  Wordperfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/289,653
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/887,284
;      FILING DATE:  May 22, 1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warren M. Cheek, Jr.
;      REGISTRATION NUMBER:  33,367
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-8850
;      TELEFAX:
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1140 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
;    HYPOTHETICAL:
;    ANTI-SENSE:
;    FRAGMENT TYPE:
;    ORIGINAL SOURCE:
;      ORGANISM:  Fusarium sp.
;      STRAIN:  S-19-5 (IFO 8884)
;      INDIVIDUAL ISOLATE:
;      DEVELOPMENTAL STAGE:
;      HAPLOTYPE:
;      TISSUE TYPE:
;      CELL TYPE:
;      CELL LINE:
;      ORGANELLE:
;    IMMEDIATE SOURCE:
;      LIBRARY:
;      CLONE:
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:
;      MAP POSITION:
;      UNITS:
;    FEATURE:
;      NAME/KEY:
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:
;    PUBLICATION INFORMATION:
;      AUTHORS:
;      TITLE:
;      JOURNAL:
;      VOLUME:
;      ISSUE:
;      PAGES:
;      DATE:
;      DOCUMENT NUMBER:
;      FILING DATE:
;      PUBLICATION DATE:
;      RELEVANT RESIDUES IN SEQ ID NO:
US-08-289-653-2


  Query Match             26.5%;  Score 26.8;  DB 1;  Length 1140;
  Best Local Similarity   57.0%;  Pred. No. 12;
  Matches   49;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy      10 ccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtcaggggaag 69
           ||||||| | ||||| | |   | | |    ||   | || |||| ||    | ||  ||
Db     124 CCTTGAGCTTGACAATGTACTTGCCAGCAATGGGCTGCGAACCACGAGCCTCGAGGAGAG 65

Qy      70 gtccccggtccaggccagctgctgag 95
           |  |  | ||||||| | |  || | |
Db      64 GAGCGGGTTCCAGGACCGGGGCCGCG 39


RESULT  12
US-08-289-653-1/c
; Sequence 1, Application US/08289653
; Patent No. 5543322
;  GENERAL INFORMATION:
;    APPLICANT:  Kazuaki KITANO et al.
;    TITLE OF INVENTION:  DNA AND ITS USE
;    NUMBER OF SEQUENCES:  4
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wenderoth, Lind & Ponack
;      STREET:  805 Fifteenth Street, N.W., #700
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 5.25 inch, 500 kb
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  Wordperfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/289,653
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
```

```
;       APPLICATION NUMBER:  US/07/887,284
;       FILING DATE:  May 22, 1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Warren M. Cheek, Jr.
;       REGISTRATION NUMBER:  33,367
;       REFERENCE/DOCKET NUMBER:
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  202-371-8850
;       TELEFAX:
;       TELEX:
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2845 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  Genomic DNA
;     HYPOTHETICAL:
;     ANTI-SENSE:
;     FRAGMENT TYPE:
;     ORIGINAL SOURCE:
;       ORGANISM:  Fusarium sp.
;       STRAIN:  S-19-5 (IFO 8884)
;       INDIVIDUAL ISOLATE:
;       DEVELOPMENTAL STAGE:
;       HAPLOTYPE:
;       TISSUE TYPE:
;       CELL TYPE:
;       CELL LINE:
;       ORGANELLE:
;     IMMEDIATE SOURCE:
;       LIBRARY:
;       CLONE:
;     POSITION IN GENOME:
;       CHROMOSOME/SEGMENT:
;       MAP POSITION:
;       UNITS:
;     FEATURE:
;       NAME/KEY:
;       LOCATION:
;       IDENTIFICATION METHOD:
;       OTHER INFORMATION:
;     PUBLICATION INFORMATION:
;       AUTHORS:
;       TITLE:
;       JOURNAL:
;       VOLUME:
;       ISSUE:
;       PAGES:
;       DATE:
;       DOCUMENT NUMBER:
;       FILING DATE:
;       PUBLICATION DATE:
;       RELEVANT RESIDUES IN SEQ ID NO:
US-08-289-653-1


  Query Match             26.5%;  Score 26.8;  DB 1;  Length 2845;
  Best Local Similarity   57.0%;  Pred. No. 14;
  Matches   49;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy      10 ccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtcaggggaag 69
           ||||||| | ||||| | |   | | |    ||   | || |||| ||    | ||  ||
Db     665 CCTTGAGCTTGACAATGTACTTGCCAGCAATGGGCTGCGAACCACGAGCCTCGAGGAGAG 606

Qy      70 gtccccggtccaggccagctgctgag 95
           |  |  | ||||| | |  || | |
Db     605 GAGCGGGTTCCAGGACCGGGGCCGCG 580


RESULT  13
US-08-164-292B-15/c
; Sequence 15, Application US/08164292B
; Patent No. 5820868
;  GENERAL INFORMATION:
;    APPLICANT:  MITTAL, SURESH K.
;    APPLICANT:  GRAHAM, FRANK L.
;    APPLICANT:  PREVEC, LUDVIK
;    APPLICANT:  BABIUK, LORNE A.
;    TITLE OF INVENTION:  RECOMBINANT PROTEIN PRODUCTION IN BOVINE
;    TITLE OF INVENTION:  ADENOVIRUS EXPRESSION VECTOR SYSTEM
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  345 California Street
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94104-2675
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/164,292B
;      FILING DATE:  09-DEC-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GRACEY, NANCY J.
;      REGISTRATION NUMBER:  28,216
;      REFERENCE/DOCKET NUMBER:  29310-20021.00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 677-7000
;      TELEFAX:  (415) 677-7522
;      TELEX:  34-0154
;  INFORMATION FOR SEQ ID NO:  15:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5100 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  2..418
US-08-164-292B-15


  Query Match             26.5%;  Score 26.8;  DB 1;  Length 5100;
  Best Local Similarity   61.4%;  Pred. No. 16;
  Matches   43;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy       9 gccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtcaggggaa 68
           | |||||| |||| || | | |||| |  |||| ||| |  |     | |  ||||
Db    1682 GTCTTGAGGAGGACGGGGAGGCACTCGGCGTGGGAACAGGAACAGCGGATGCTGCCGGAA 1623

Qy      69 ggtccccggt 78
           | |||| |||
Db    1622 GCTCCCTGGT 1613


RESULT  14
US-08-164-292B-17/c
; Sequence 17, Application US/08164292B
; Patent No. 5820868
;  GENERAL INFORMATION:
;    APPLICANT:  MITTAL, SURESH K.
;    APPLICANT:  GRAHAM, FRANK L.
;    APPLICANT:  PREVEC, LUDVIK
;    APPLICANT:  BABIUK, LORNE A.
;    TITLE OF INVENTION:  RECOMBINANT PROTEIN PRODUCTION IN BOVINE
;    TITLE OF INVENTION:  ADENOVIRUS EXPRESSION VECTOR SYSTEM
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
```

```
;       STREET:  345 California Street
;       CITY:  San Francisco
;       STATE:  California
;       COUNTRY:  USA
;       ZIP:  94104-2675
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/164,292B
;       FILING DATE:  09-DEC-1993
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  GRACEY, NANCY J.
;       REGISTRATION NUMBER:  28,216
;       REFERENCE/DOCKET NUMBER:  29310-20021.00
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (415) 677-7000
;       TELEFAX:  (415) 677-7522
;       TELEX:  34-0154
;   INFORMATION FOR SEQ ID NO:  17:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  5100 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  408..1331
US-08-164-292B-17


  Query Match             26.5%;  Score 26.8;  DB 1;  Length 5100;
  Best Local Similarity   61.4%;  Pred. No. 16;
  Matches   43;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy       9 gccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtcaggggaa 68
           | ||||||  ||||  ||  | | |||| |  ||||  ||| |  |     | |  ||||
Db    1682 GTCTTGAGGAGGACGGGGAGGCACTCGGCGTGGGAACAGGAACAGCGGATGCTGCCGGAA 1623

Qy      69 ggtccccggt 78
           | ||||  |||
Db    1622 GCTCCCTGGT 1613


RESULT  15
US-08-164-292B-19/c
; Sequence 19, Application US/08164292B
; Patent No. 5820868
;  GENERAL INFORMATION:
;    APPLICANT:  MITTAL, SURESH K.
;    APPLICANT:  GRAHAM, FRANK L.
;    APPLICANT:  PREVEC, LUDVIK
;    APPLICANT:  BABIUK, LORNE A.
;    TITLE OF INVENTION:  RECOMBINANT PROTEIN PRODUCTION IN BOVINE
;    TITLE OF INVENTION:  ADENOVIRUS EXPRESSION VECTOR SYSTEM
;    NUMBER OF SEQUENCES:  34
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  345 California Street
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94104-2675
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/164,292B
;      FILING DATE:  09-DEC-1993
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  GRACEY, NANCY J.
;      REGISTRATION NUMBER:  28,216
;      REFERENCE/DOCKET NUMBER:  29310-20021.00
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 677-7000
;      TELEFAX:  (415) 677-7522
;      TELEX:  34-0154
;  INFORMATION FOR SEQ ID NO:  19:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5100 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  529..954
US-08-164-292B-19


  Query Match             26.5%;  Score 26.8;  DB 1;  Length 5100;
  Best Local Similarity   61.4%;  Pred. No. 16;
  Matches   43;  Conservative    0;  Mismatches   27;  Indels    0;  Gaps    0;

Qy       9 gccttgagatggacaaggcagaagtcggtggaggaaggggagccaccaggtcaggggaa 68
           | ||||||  ||||  ||  | | |||| |  ||||  ||| |  |     | |  ||||
Db    1682 GTCTTGAGGAGGACGGGGAGGCACTCGGCGTGGGAACAGGAACAGCGGATGCTGCCGGAA 1623

Qy      69 ggtccccggt 78
           | ||||  |||
Db    1622 GCTCCCTGGT 1613


Search completed:  October  3, 2002, 16:22:14
Job time:  16974 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 14:49:24 ; Search time 7316.32 Seconds
(without alignments)
288.886 Million cell updates/sec

Title: US-09-826-581-5_COPY_500_600
Perfect score: 101
Sequence: 1 ctgtccccgcaggccccatt.....atgcaggagcacacctgcta 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	101	100.0	989	6	AX281579	AX281579 Sequence
2	101	100.0	1647	6	AX281582	AX281582 Sequence
3	101	100.0	2109	6	AX099776	AX099776 Sequence
4	101	100.0	2115	6	AX099802	AX099802 Sequence
5	101	100.0	2115	9	AF214519	AF214519 Homo sapi
c 6	101	100.0	152129	2	AC027416	AC027416 Homo sapi
c 7	101	100.0	196554	2	AC073128	AC073128 Homo sapi
8	97.8	96.8	2290	9	HSA249977	AJ249977 Homo sapi
c 9	89.6	88.7	206854	9	AC009974	AC009974 Homo sapi
10	72.4	71.7	1867	6	AX099774	AX099774 Sequence
11	72.4	71.7	1873	4	AF214520	AF214520 Sus scrof
12	72.4	71.7	1873	6	AX099800	AX099800 Sequence
13	72.4	71.7	2022	6	AX099804	AX099804 Sequence
14	72.4	71.7	5888	4	AF214521	AF214521 Sus scrof
c 15	72.4	71.7	227724	2	AF336381	AF336381 Mus muscu
16	33.4	33.1	3503	9	HSL81675	L81675 Homo sapien
17	33.4	33.1	81811	2	AC063981	AC063981 Homo sapi
c 18	33.4	33.1	82237	9	AC003677	AC003677 Human Pl
19	33.4	33.1	108123	9	AC034206	AC034206 Homo sapi
20	33.4	33.1	159420	2	AC009017	AC009017 Homo sapi
21	32	31.7	790	33	AC076924	Ac076924 Giardia i
c 22	32	31.7	975	33	AC078451	Ac078451 Giardia i
23	32	31.7	988	33	AC028281	Ac028281 Giardia i
24	32	31.7	1927	3	AF071896	AF071896 Giardia i
c 25	31.2	30.9	17570	1	AE007063	AE007063 Mycobacte
c 26	31.2	30.9	39430	1	MTCY49	Z73966 Mycobacteri
c 27	30.4	30.1	736	33	AC053985	Ac053985 Giardia i
c 28	30.4	30.1	804	33	AC053984	Ac053984 Giardia i
29	30.4	30.1	967	33	AC089076	Ac089076 Giardia i
30	30.4	30.1	1283	3	AF050755	AF050755 Giardia i
31	30.4	30.1	1306	3	AF050754	AF050754 Giardia i
32	30.4	30.1	3989	3	AF071897	AF071897 Giardia i
33	30.4	30.1	5791	1	AF117694	AF117694 Rhizobium
c 34	30.2	29.9	11085	1	AE004450	AE004450 Pseudomon
35	29.6	29.3	3603	1	AF035395	AF035395 Pseudomon
36	29.6	29.3	38810	3	AC091120	AC091120 Leishmani
37	29.4	29.1	850	33	AC042098	Ac042098 Giardia i
c 38	29.4	29.1	214609	2	AC024651	AC024651 Homo sapi
c 39	29.2	28.9	56425	7	AF165214	AF165214 Bacteriop
40	29	28.7	1575	3	AY069691	AY069691 Drosophil
41	29	28.7	69061	2	AC012986	AC012986 Drosophil
42	29	28.7	158628	2	AC020653	AC020653 Homo sapi
43	29	28.7	168469	3	AC007886	AC007886 Drosophil
c 44	29	28.7	173729	2	AC012236	AC012236 Homo sapi
45	29	28.7	228448	3	AE003772	AE003772 Drosophil

ALIGNMENTS

RESULT 1
AX281579
LOCUS AX281579 989 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 2 from Patent WO0177305.
ACCESSION AX281579
VERSION AX281579.1 GI:16608830
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Andersson,L., Luthman,H. and Marklund,S.
TITLE Variants of the human amp-activated protein kinase gamma 3 subunit
JOURNAL Patent: WO 0177305-A 2 18-OCT-2001;
Arexis AB (SE)
FEATURES Location/Qualifiers
source 1. .989
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 229 a 306 c 286 g 168 t
ORIGIN

```
  Query Match            100.0%;  Score 101;  DB 6;  Length 989;
  Best Local Similarity  100.0%;  Pred. No. 1.5e-17;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    793 CTGTCCCCGCAGGCCCCATTTCCCAAGCTGGGCTGGGATGACGAACTGCGGAAACCCGGC 852

Qy     61 gcccagatctacatgcgcttcatgcaggagcacacctgcta 101
          |||||||||||||||||||||||||||||||||||||||||
Db    853 GCCCAGATCTACATGCGCTTCATGCAGGAGCACACCTGCTA 893


RESULT   2
AX281582
LOCUS       AX281582                1647 bp    DNA     linear   PAT 02-NOV-2001
DEFINITION  Sequence 5 from Patent WO0177305.
ACCESSION   AX281582
VERSION     AX281582.1  GI:16608833
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Andersson,L., Luthman,H. and Marklund,S.
  TITLE     Variants of the human amp-activated protein kinase gamma 3 subunit
  JOURNAL   Patent: WO 0177305-A 5 18-OCT-2001;
            Arexis AB (SE)
FEATURES             Location/Qualifiers
     source          1. .1647
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             20. .1489
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD10589.1"
                     /db_xref="GI:16608834"
                     /translation="MEPGLEHALRRTPSWSSLGGSEHQEMSFLEQENSSSWPSPAVTS
                     SSERIRGKRRAKALRWTRQKSVEEGEPPGQGEGPRSRPAAESTGLEATFPKTTPLAQA
                     DPAGVGTPPTGWDCLPSDCTASAAGSSTDDVELATEFPATEAWECELEGLLEERPALC
                     LSPQAPFPKLGWDDEILRKPGAQIYMRFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFA
                     LVANGVRAAPLWDSKKQSFVGMLTITDFILVLHRYYRSPLVQIYEIEQHKIETWREIY
                     LQGCFKPLVSISPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIF
                     GSLLPRPSFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSALPVVNECGQV
                     VGLYSRFDVIHLAAQQTYNHLDMSVGEALRQRTLCLEGVLSCQPHESLGEVIDRIARE
                     QVHRLVLVDETQHLLGVVSLSDILQALVLSPAGIDALGA"
BASE COUNT      346 a    502 c    462 g    337 t
ORIGIN


  Query Match            100.0%;  Score 101;  DB 6;  Length 1647;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-17;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    500 CTGTCCCCGCAGGCCCCATTTCCCAAGCTGGGCTGGGATGACGAACTGCGGAAACCCGGC 559

Qy     61 gcccagatctacatgcgcttcatgcaggagcacacctgcta 101
          |||||||||||||||||||||||||||||||||||||||||
Db    560 GCCCAGATCTACATGCGCTTCATGCAGGAGCACACCTGCTA 600


RESULT   3
AX099776
LOCUS       AX099776                2109 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 3 from Patent WO0120003.
ACCESSION   AX099776
VERSION     AX099776.1  GI:13538810
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2109)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 3 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .2109
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             472. .1389
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35799.1"
                     /db_xref="GI:13538811"
                     /translation="MRFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAP
                     LWDSKKQSFVGMLTITDFILVLHRYYRSPLVQIYEIEQHKIETWREIYLQGCFKPLVS
                     ISPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLPRPSFL
                     YRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSALPVVNECGQVVGLYSRFDVI
                     HLAAQQTYNHLDMSVGEALRQRTLCLEGVLSCQPHESLGEVIDRIAREQVHRLVLVDE
                     TQHLLGVVSLSDILQALVLSPAGIDALGA"
BASE COUNT      458 a    621 c    560 g    470 t
ORIGIN


  Query Match            100.0%;  Score 101;  DB 6;  Length 2109;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-17;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 ctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggc 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    400 CTGTCCCCGCAGGCCCCATTTCCCAAGCTGGGCTGGGATGACGAACTGCGGAAACCCGGC 459

Qy     61 gcccagatctacatgcgcttcatgcaggagcacacctgcta 101
          |||||||||||||||||||||||||||||||||||||||||
Db    460 GCCCAGATCTACATGCGCTTCATGCAGGAGCACACCTGCTA 500


RESULT   4
AX099802
LOCUS       AX099802                2115 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 29 from Patent WO0120003.
ACCESSION   AX099802
VERSION     AX099802.1  GI:13538836
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2115)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 29 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .2115
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             1. .1395
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35801.1"
                     /db_xref="GI:13538837"
```

```
                     /translation="MSFLEQENSSSWPSPAVTSSSERIRGKRRAKALRWTRQKSVEEG
                     EPPGQGEGPRSRPTAESTGLEATFPKTTPLAQADPAGVGTPPTGWDCLPSDCTASAAG
                     SSTDDVELATEFPATEAWECELEGLLEERPALCLSPQAPFPKLGWDDELRKPGAQIYM
                     RFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEQHKIETWREIYLQGCFKPLVSISPNDSLFEAVYTLI
                     KNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNECGQVVGLYSRFDVIHLAAQQTYNHLDMSV
                     GEALRQRTLCLEGVLSCQPHESLGEVIDRIAREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT       460 a    622 c    562 g    471 t
ORIGIN


  Query Match           100.0%;  Score 101;  DB 6;  Length 2115;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-17;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     406 CTGTCCCCGCAGGCCCCATTTCCCAAGCTGGGCTGGGATGACGAACTGCGGAAACCCGGC 465

Qy      61 gcccagatctacatgcgcttcatgcaggagcacacctgcta 101
           |||||||||||||||||||||||||||||||||||||||||
Db     466 GCCCAGATCTACATGCGCTTCATGCAGGAGCACACCTGCTA 506


RESULT   5
AF214519
LOCUS       AF214519                2115 bp    mRNA    linear   PRI 03-JUN-2000
DEFINITION  Homo sapiens AMP-activated protein kinase gamma subunit (PRKAG3)
            mRNA, complete cds.
ACCESSION   AF214519
VERSION     AF214519.1  GI:8215681
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2115)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A., Thelander,M.,
            Rogel-Gaillard,C., Paul,S., Iannuccelli,N., Rask,L., Ronne,H.,
            Lundstrom,K., Reinsch,N., Gellin,J., Kalm,E., Roy,P.L., Chardon,P.
            and Andersson,L.
  TITLE     A mutation in PRKAG3 associated with excess glycogen content in pig
            skeletal muscle
  JOURNAL   Science 288 (5469), 1248-1251 (2000)
  MEDLINE   20280150
   PUBMED   10818001
REFERENCE   2  (bases 1 to 2115)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A.,
            Rogel-Gaillard,C., Paul,S., Gellin,J., Lundstrom,K., Reinsch,N.,
            Kalm,E., Le Roy,P., Chardon,P. and Andersson,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-1999) Dept Animal Breeding and Genetics, Swedish
            University of Agricultural Sciences, BMC box 597, Uppsala 751 24,
            Sweden
FEATURES             Location/Qualifiers
     source          1. .2115
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2p"
                     /tissue_type="skeletal muscle"
     gene            1. .2115
                     /gene="PRKAG3"
     CDS             1. .1395
                     /gene="PRKAG3"
                     /note="AMPKG3"
                     /codon_start=1
                     /product="AMP-activated protein kinase gamma subunit"
                     /protein_id="AAF73987.1"
                     /db_xref="GI:8215682"
                     /translation="MSFLEQENSSSWPSPAVTSSSERIRGKRRAKALRWTRQKSVEEG
                     EPPGQGEGPRSRPTAESTGLEATFPKTTPLAQADPAGVGTPPTGWDCLPSDCTASAAG
                     SSTDDVELATEFPATEAWECELEGLLEERPALCLSPQAPFPKLGWDDELRKPGAQIYM
                     RFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEQHKIETWREIYLQGCFKPLVSISPNDSLFEAVYTLI
                     KNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNECGQVVGLYSRFDVIHLAAQQTYNHLDMSV
                     GEALRQRTLCLEGVLSCQPHESLGEVIDRIAREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT       460 a    622 c    562 g    471 t
ORIGIN


  Query Match           100.0%;  Score 101;  DB 9;  Length 2115;
  Best Local Similarity  100.0%;  Pred. No. 1.4e-17;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     406 CTGTCCCCGCAGGCCCCATTTCCCAAGCTGGGCTGGGATGACGAACTGCGGAAACCCGGC 465

Qy      61 gcccagatctacatgcgcttcatgcaggagcacacctgcta 101
           |||||||||||||||||||||||||||||||||||||||||
Db     466 GCCCAGATCTACATGCGCTTCATGCAGGAGCACACCTGCTA 506


RESULT   6
AC027416/c
LOCUS       AC027416              152129 bp    DNA     linear   HTG 07-JUN-2000
DEFINITION  Homo sapiens clone RP11-504G11, WORKING DRAFT SEQUENCE, 32
            unordered pieces.
ACCESSION   AC027416
VERSION     AC027416.2  GI:8317289
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 152129)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-504G11
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 152129)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jun 7, 2000 this sequence version replaced gi:7342115.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ---------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
```

```
    Web site: http://www-seq.wi.mit.edu
    Contact: sequence_submissions@genome.wi.mit.edu
-------------- Project Information
    Center project name: L7458
    Center clone name: 504_G_11
-------------- Summary Statistics
    Sequencing vector: M13; M77815; 100% of reads
    Chemistry: Dye-terminator Big Dye; 100% of reads
    Assembly program: Phrap; version 0.960731
    Consensus quality: 135376 bases at least Q40
    Consensus quality: 143264 bases at least Q30
    Consensus quality: 146503 bases at least Q20
    Insert size: 161000; agarose-fp
    Insert size: 149029; sum-of-contigs
    Quality coverage: 3.1 in Q20 bases; agarose-fp
    Quality coverage: 3.3 in Q20 bases; sum-of-contigs
--------------.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 32 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*        1     1005: contig of 1005 bp in length
*     1006 1105: gap of      100 bp
*     1106     2402: contig of 1297 bp in length
*     2403 2502: gap of      100 bp
*     2503     3823: contig of 1321 bp in length
*     3824 3923: gap of      100 bp
*     3924     5020: contig of 1097 bp in length
*     5021 5120: gap of      100 bp
*     5121     6161: contig of 1041 bp in length
*     6162 6261: gap of      100 bp
*     6262     7547: contig of 1286 bp in length
*     7548 7647: gap of      100 bp
*     7648     9983: contig of 2336 bp in length
*     9984 10083: gap of      100 bp
*    10084    12556: contig of 2473 bp in length
*    12557 12656: gap of      100 bp
*    12657    15043: contig of 2387 bp in length
*    15044 15143: gap of      100 bp
*    15144    17123: contig of 1980 bp in length
*    17124 17223: gap of      100 bp
*    17224    19466: contig of 2243 bp in length
*    19467 19566: gap of      100 bp
*    19567    21928: contig of 2362 bp in length
*    21929 22028: gap of      100 bp
*    22029    24319: contig of 2291 bp in length
*    24320 24419: gap of      100 bp
*    24420    27059: contig of 2640 bp in length
*    27060 27159: gap of      100 bp
*    27160    30170: contig of 3011 bp in length
*    30171 30270: gap of      100 bp
*    30271    33968: contig of 3698 bp in length
*    33969 34068: gap of      100 bp
*    34069    38179: contig of 4111 bp in length
*    38180 38279: gap of      100 bp
*    38280    42366: contig of 4087 bp in length
*    42367 42466: gap of      100 bp
*    42467    46365: contig of 3899 bp in length
*    46366 46465: gap of      100 bp
*    46466    51285: contig of 4820 bp in length
*    51286 51385: gap of      100 bp
*    51386    55871: contig of 4486 bp in length
*    55872 55971: gap of      100 bp
*    55972    60595: contig of 4624 bp in length
*    60596 60695: gap of      100 bp
*    60696    66595: contig of 5900 bp in length
*    66596 66695: gap of      100 bp
*    66696    73218: contig of 6523 bp in length
*    73219 73318: gap of      100 bp
*    73319    77115: contig of 3797 bp in length
*    77116 77215: gap of      100 bp
*    77216    85022: contig of 7807 bp in length
*    85023 85122: gap of      100 bp
*    85123    93314: contig of 8192 bp in length
*    93315 93414: gap of      100 bp
*    93415   101193: contig of 7779 bp in length
*   101194 101293: gap of      100 bp
*   101294   113090: contig of 11797 bp in length
*   113091 113190: gap of      100 bp
*   113191   123496: contig of 10306 bp in length
*   123497 123596: gap of      100 bp
*   123597   137837: contig of 14241 bp in length
*   137838 137937: gap of      100 bp
*   137938   152129: contig of 14192 bp in length.
FEATURES             Location/Qualifiers
     source          1. .152129
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-504G11"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1005
                     /note="assembly_fragment"
     misc_feature    1106. .2402
                     /note="assembly_fragment"
     misc_feature    2503. .3823
                     /note="assembly_fragment"
     misc_feature    3924. .5020
                     /note="assembly_fragment"
     misc_feature    5121. .6161
                     /note="assembly_fragment"
     misc_feature    6262. .7547
                     /note="assembly_fragment"
     misc_feature    7648. .9983
                     /note="assembly_fragment"
     misc_feature    10084. .12556
                     /note="assembly_fragment"
     misc_feature    12657. .15043
                     /note="assembly_fragment"
     misc_feature    15144. .17123
                     /note="assembly_fragment"
     misc_feature    17224. .19466
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    19567. .21928
                     /note="assembly_fragment"
     misc_feature    22029. .24319
                     /note="assembly_fragment"
     misc_feature    24420. .27059
                     /note="assembly_fragment"
     misc_feature    27160. .30170
                     /note="assembly_fragment"
     misc_feature    30271. .33968
                     /note="assembly_fragment"
     misc_feature    34069. .38179
                     /note="assembly_fragment"
     misc_feature    38280. .42366
                     /note="assembly_fragment"
     misc_feature    42467. .46365
                     /note="assembly_fragment"
     misc_feature    46466. .51285
                     /note="assembly_fragment"
     misc_feature    51386. .55871
                     /note="assembly_fragment"
     misc_feature    55972. .60595
                     /note="assembly_fragment"
     misc_feature    60696. .66595
                     /note="assembly_fragment"
     misc_feature    66696. .73218
                     /note="assembly_fragment"
     misc_feature    73319. .77115
                     /note="assembly_fragment"
```

```
     misc_feature    77216. .85022
                     /note="assembly_fragment"
     misc_feature    85123. .93314
                     /note="assembly_fragment"
     misc_feature    93415. .101193
                     /note="assembly_fragment"
     misc_feature    101294. .113090
                     /note="assembly_fragment"

  Query Match             100.0%;  Score 101;  DB 2;  Length 152129;
  Best Local Similarity   100.0%;  Pred. No. 1e-17;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   38135 CTGTCCCCGCAGGCCCCATTTCCCAAGCTGGGCTGGGATGACGAACTGCGGAAACCCGGC 38076

Qy      61 gcccagatctacatgcgcttcatgcaggagcacacctgcta 101
           |||||||||||||||||||||||||||||||||||||||||
Db   38075 GCCCAGATCTACATGCGCTTCATGCAGGAGCACACCTGCTA 38035


RESULT   7
AC073128/c
LOCUS       AC073128              196554 bp    DNA     linear   HTG 21-FEB-2001
DEFINITION  Homo sapiens chromosome 2 clone RP11-647O5, WORKING DRAFT SEQUENCE,
            17 unordered pieces.
ACCESSION   AC073128
VERSION     AC073128.3  GI:13027579
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 196554)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 196554)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JUN-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Feb 21, 2001 this sequence version replaced gi:8469048.

            ---------------- Genome Center ----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: H_NH0647O05
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 98%
            Sequencing vector: plasmid; 0%
            Chemistry: Dye-primer ET; 98% of reads
            Chemistry: Dye-terminator Big Dye; 0% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 187795 bases at least Q40
            Consensus quality: 190513 bases at least Q30
            Consensus quality: 192099 bases at least Q20
            Insert size: 200000; agarose-fp
            Insert size: 194954; sum-of-contigs
            Quality coverage: 5.58 in Q20 bases; agarose-fp
            Quality coverage: 5.67 in Q20 bases; sum-of-contigs
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 17 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     1157: contig of 1157 bp in length
            *    1158     1257: gap of unknown length
            *    1258     3600: contig of 2343 bp in length
            *    3601     3700: gap of unknown length
            *    3701     5103: contig of 1403 bp in length
            *    5104     5203: gap of unknown length
            *    5204     8524: contig of 3321 bp in length
            *    8525     8624: gap of unknown length
            *    8625    11856: contig of 3232 bp in length
            *   11857    11956: gap of unknown length
            *   11957    15783: contig of 3827 bp in length
            *   15784    15883: gap of unknown length
            *   15884    21906: contig of 6023 bp in length
            *   21907    22006: gap of unknown length
            *   22007    28887: contig of 6881 bp in length
            *   28888    28987: gap of unknown length
            *   28988    35255: contig of 6268 bp in length
            *   35256    35355: gap of unknown length
            *   35356    44642: contig of 9287 bp in length
            *   44643    44742: gap of unknown length
            *   44743    58275: contig of 13533 bp in length
            *   58276    58375: gap of unknown length
            *   58376    73816: contig of 15441 bp in length
            *   73817    73916: gap of unknown length
            *   73917    92140: contig of 18224 bp in length
            *   92141    92240: gap of unknown length
            *   92241   113337: contig of 21097 bp in length
            *  113338   113437: gap of unknown length
            *  113438   130325: contig of 16888 bp in length
            *  130326   130425: gap of unknown length
            *  130426   149287: contig of 18862 bp in length
            *  149288   149387: gap of unknown length
            *  149388   196554: contig of 47167 bp in length.
FEATURES             Location/Qualifiers
     source          1. .196554
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /clone="RP11-647O5"
     misc_feature    1. .1157
                     /note="assembly_name:Contig17"
     misc_feature    1258. .3600
                     /note="assembly_name:Contig18"
     misc_feature    3701. .5103
                     /note="assembly_name:Contig19
                     clone_end:T7
                     vector_side:right"
     misc_feature    5204. .8524
                     /note="assembly_name:Contig20"
     misc_feature    8625. .11856
                     /note="assembly_name:Contig21"
     misc_feature    11957. .15783
                     /note="assembly_name:Contig22"
     misc_feature    15884. .21906
                     /note="assembly_name:Contig23"
     misc_feature    22007. .28887
                     /note="assembly_name:Contig24"
     misc_feature    28988. .35255
                     /note="assembly_name:Contig25"
     misc_feature    35356. .44642
                     /note="assembly_name:Contig26"
     misc_feature    44743. .58275
                     /note="assembly_name:Contig27"
     misc_feature    58376. .73816
                     /note="assembly_name:Contig28"
     misc_feature    73917. .92140
                     /note="assembly_name:Contig29"
     misc_feature    92241. .113337
                     /note="assembly_name:Contig30"
     misc_feature    113438. .130325
                     /note="assembly_name:Contig31"
     misc_feature    130426. .149287
```

```
                     /note="assembly_name:Contig32"
     misc_feature    149388. .196554
                     /note="assembly_name:Contig33
                     clone_end:SP6
                     vector_side:right"
BASE COUNT     52296 a   46993 c   45889 g   49770 t    1606 others
ORIGIN


  Query Match             100.0%;  Score 101;  DB 2;  Length 196554;
  Best Local Similarity   100.0%;  Pred. No. 1e-17;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   62634 CTGTCCCCGCAGGCCCCATTTCCCAAGCTGGGCTGGGATGACGAACTGCGGAAACCCGGC 62575

Qy      61 gcccagatctacatgcgcttcatgcaggagcacacctgcta 101
           |||||||||||||||||||||||||||||||||||||||||
Db   62574 GCCCAGATCTACATGCGCTTCATGCAGGAGCACACCTGCTA 62534


RESULT   8
HSA249977
LOCUS       HSA249977               2290 bp    mRNA    linear   PRI 07-APR-2000
DEFINITION  Homo sapiens mRNA for AMP-activated protein kinase gamma 3 subunit
            (AMPK gamma 3 gene).
ACCESSION   AJ249977
VERSION     AJ249977.1  GI:6688200
KEYWORDS    AMP-activated protein kinase; AMPK gamma 3 gene; gamma 3 subunit.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2290)
  AUTHORS   Cheung,P.C., Salt,I.P., Davies,S.P., Hardie,D.G. and Carling,D.
  TITLE     Characterization of AMP-activated protein kinase gamma-subunit
            isoforms and their role in AMP binding
  JOURNAL   Biochem. J. 346 Pt 3, 659-669 (2000)
  MEDLINE   20164049
REFERENCE   2  (bases 1 to 2290)
  AUTHORS   Carling,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-OCT-1999) Carling D., Cellular Stress Group, MRC
            Clinical Sciences Centre, Hammersmith Hospital, DuCane Road,
            London, W12 0NN, UNITED KINGDOM
FEATURES             Location/Qualifiers
     source          1. .2290
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            22. .1500
                     /gene="AMPK gamma 3"
     CDS             22. .1500
                     /gene="AMPK gamma 3"
                     /function="AMP-activated protein kinase regulatory
                     subunit"
                     /codon_start=1
                     /evidence=experimental
                     /product="AMP-activated protein kinase gamma 3 subunit"
                     /protein_id="CAB65117.1"
                     /db_xref="GI:6688201"
                     /translation="MEPGLEHALRRTPSWSSLGGSEHQEMSFLEQENSSSWPSPAVTS
                     SSERIRGKRRAKALRWTRQKSVEEGEPPGQGEGPRSRPAAESTGLEATFPKTTPLAQA
                     DPAGVGTPPTGWDCLPSDCTASAAGSSTDDVELATEFPATEAWECELEGLLEERPALC
                     LSPQAPFPKLGWDDELRKPGAQIYMRFIEEHTCYDAMATSSKLVIFDTMLEIKKAFFA
                     LVANGVRAAPLWDSKKQSFVGMLTITDFILVLHRYYRSPLVQIYEIEQHKIETWREIY
                     LQGCFKPLVSISPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIF
                     GSLLPRPSFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSALPVVNECGQV
                     VGLYSRFDVIHLAAQQTYNHLDMSVGEALRKRTLCLEGVLSCQPHESLGEVIDRIARE
                     QVHRLVLVDETQHLLGVVSLSDILQALVLSPAGIDPSGPEKI"
BASE COUNT     501 a     674 c     617 g     498 t
ORIGIN
```

```
  Query Match             96.8%;  Score 97.8;  DB 9;  Length 2290;
  Best Local Similarity   98.0%;  Pred. No. 1e-16;
  Matches   99;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy       1 ctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     502 CTGTCCCCGCAGGCCCCATTTCCCAAGCTGGGCTGGGATGACGAACTGCGGAAACCCGGC 561

Qy      61 gcccagatctacatgcgcttcatgcaggagcacacctgcta 101
           |||||||||||||||||||||||  ||||||||||||||||
Db     562 GCCCAGATCTACATGCGCTTCATCGAGGAGCACACCTGCTA 602


RESULT   9
AC009974/c
LOCUS       AC009974               206854 bp    DNA     linear   PRI 09-JAN-2002
DEFINITION  Homo sapiens BAC clone RP11-459I19 from 2, complete sequence.
ACCESSION   AC009974
VERSION     AC009974.9  GI:16799058
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 206854)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
REFERENCE   2  (bases 1 to 206854)
  AUTHORS   Harris,A. and Cotton,M.
  TITLE     The sequence of Homo sapiens BAC clone RP11-459I19
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 206854)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 206854)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-NOV-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 206854)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JAN-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   6  (bases 1 to 206854)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JAN-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Nov 8, 2001 this sequence version replaced gi:13431203.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0459I19
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
```

all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E., Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (http://bacpac.med.buffalo.edu)
VECTOR: pBACe3.6
NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-1077K22; the clone sequenced to the right is RP11-64705. Actual start of this clone is at base position 1 of RP11-459I19; actual end is at base position 206854 of RP11-459I19.

Data from AC079810 and AC073128 was used to finish this clone, AC009974. Polymorphisms have been identified between AC073128 and AC009974. A single plasmid region exists beteen 38812-38903. Am unresolved tandem in the HERV SVA exists between 184390-185163. PCR suggests that approximately 1700 bps are missing.

FEATURES Location/Qualifiers
source 1. .206854
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-459I19"
/clone_lib="RPCI-11"
misc_feature 1. .37
/note="match to EST BE314060 (NID:g9134719)"
misc_feature 1. .37
/note="match to EST BI059713 (NID:g14467240)"
misc_feature 1. .37
/note="match to EST BF183086 (NID:g11061273)"
misc_feature 1. .37
/note="match to EST AL567345 (NID:g12920610)"
misc_feature 1. .37
/note="match to EST AW880850 (NID:g8042860)"
misc_feature 1. .37
/note="match to EST BF304755 (NID:g11251653)"
misc_feature 1. .37
/note="similar to Homo sapiens EST BI114348 (NID:g14565249)"
misc_feature 1. .37
/note="match to EST BG477625 (NID:g13409904)"
misc_feature 1. .37
/note="match to EST BE047599 (NID:g8364652) tz39c01.y1"
misc_feature 3. .37
/note="match to EST BE908408 (NID:g10402954)"
misc_feature 4. .37
/note="match to EST AI670836 (NID:g4850567) wa04g10.x1"
misc_feature 164. .662
/note="match to EST C05773 (NID:g1502549)"
misc_feature 220. .221
/note="match to EST BG470047 (NID:g13402322)"
misc_feature 281. .344
/note="match to EST BE047599 (NID:g8364652) tz39c01.y1"
misc_feature 281. .344
/note="match to EST BE908408 (NID:g10402954)"
misc_feature 281. .344
/note="match to EST BF304755 (NID:g11251653)"
misc_feature 281. .344
/note="match to EST BF183086 (NID:g11061273)"
misc_feature 281. .344
/note="similar to Homo sapiens EST BI114348 (NID:g14565249)"
misc_feature 281. .344
/note="match to EST BG477625 (NID:g13409904)"
misc_feature 281. .344
/note="match to EST BI059713 (NID:g14467240)"
misc_feature 281. .344
/note="match to EST AL567345 (NID:g12920610)"
misc_feature 281. .344
/note="match to EST BE314060 (NID:g9134719)"
misc_feature 281. .344
/note="match to EST AI670836 (NID:g4850567) wa04g10.x1"
misc_feature 281. .344
/note="match to EST AW880850 (NID:g8042860)"
misc_feature 294. .344
/note="match to EST BG470047 (NID:g13402322)"
misc_feature complement(557. .558)
/note="match to EST AA043371 (NID:g1521226) zk53e10.r1"
misc_feature 588. .929
/note="match to EST T64073 (NID:g667938) yc05d12.r1"
misc_feature 594. .763
/note="match to EST BF304755 (NID:g11251653)"
misc_feature 594. .763
/note="match to EST BG470047 (NID:g13402322)"
misc_feature 594. .763
/note="match to EST AW880850 (NID:g8042860)"
misc_feature 594. .731
/note="match to EST BE314060 (NID:g9134719)"
misc_feature 594. .764
/note="match to EST BI059713 (NID:g14467240)"
misc_feature 594. .763
/note="match to EST BG477625 (NID:g13409904)"
misc_feature 594. .742
/note="match to EST BE047599 (NID:g8364652) tz39c01.y1"
misc_feature 594. .763
/note="match to EST BE908408 (NID:g10402954)"
misc_feature 594. .763
/note="match to EST AA043371 (NID:g1521226) zk53e10.r1"
misc_feature 594. .763
/note="match to EST AI670836 (NID:g4850567) wa04g10.x1"
misc_feature 594. .763
/note="match to EST BF183086 (NID:g11061273)"
misc_feature 594. .763
/note="similar to Homo sapiens EST BI114348 (NID:g14565249)"
misc_feature 599. .763
/note="match to EST AL567345 (NID:g12920610)"
misc_feature 617. .1084
/note="match to EST AA481361 (NID:g2210913) zv44e01.r1"
misc_feature 622. .763
/note="match to EST AI860958 (NID:g5514574) wl56f05.x1"
misc_feature 622. .763
/note="similar to Mus musculus EST AI196847 (NID:g3749453) ui67e01.x1"
misc_feature 684. .763
/note="match to EST BG992568 (NID:g14396638)"
misc_feature 962. .1084
/note="match to EST AI656812 (NID:g4740791) tt54b06.x1"
misc_feature 967. .1084
/note="match to EST BE908408 (NID:g10402954)"
misc_feature 967. .1085
/note="match to EST BF304755 (NID:g11251653)"
misc_feature 967. .1091
/note="match to EST AA043371 (NID:g1521226) zk53e10.r1"
misc_feature 967. .1071

Query Match 88.7%; Score 89.6; DB 9; Length 206854;
Best Local Similarity 95.8%; Pred. No. 1.2e-14;

```
  Matches   92;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      6 cccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcccca 65
          |||    ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db 168209 CCCCGCAGCCCCATTTCCCAAGCTGGGCTGGGATGACGAACTGCGGAAACCCGGCGCCCA 168150

Qy     66 gatctacatgcgcttcatgcaggagcacacctgcta 101
          |||||||||||||||||||||||||||||||||||
Db 168149 GATCTACATGCGCTTCATGCAGGAGCACACCTGCTA 168114


RESULT  10
AX099774
LOCUS       AX099774               1867 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 1 from Patent WO0120003.
ACCESSION   AX099774
VERSION     AX099774.1  GI:13538808
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 1867)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 1 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .1867
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
     CDS             472. .1389
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35798.1"
                     /db_xref="GI:13538809"
                     /translation="MHFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAP
                     LWDSKKQSFVGMLTITDFILVLHRYYRSPLVQIYEIEEHKIETWREIYLQGCFKPLVS
                     ISPNDSLFEAVYALIKNRIHRLPVLDPVSGAVLHILTHKRLLKFLHIFGTLLPRPSFL
                     YRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSALPVVNETGQVVGLYSRFDVI
                     HLAAQQTYNHLDMNVGEALRQRTLCLEGVLSCQPHETLGEVIDRIVREQVHRLVLVDE
                     TQHLLGVVSLSDILQALVLSPAGIDALGA"
BASE COUNT      380 a    583 c    529 g    375 t
ORIGIN


  Query Match           71.7%;  Score 72.4;  DB 6;  Length 1867;
  Best Local Similarity  83.7%;  Pred. No. 8.2e-10;
  Matches   82;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      4 tccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcc 63
          |||||  |||  |  || |||| ||||||||||||||| || |||| ||| || || |||
Db    403 TCCCCAGAGGTGCTGTTACCCAGGCTGGGCTGGGATGATGAGCTGCAGAAGCCGGGGGCC 462

Qy     64 cagatctacatgcgcttcatgcaggagcacacctgcta 101
          ||| |||||||| |||||||||||||||||||||||||
Db    463 CAGGTCTACATGCACTTCATGCAGGAGCACACCTGCTA 500


RESULT  11
AF214520
LOCUS       AF214520               1873 bp    mRNA    linear   MAM 03-JUN-2000
DEFINITION  Sus scrofa AMP-activated protein kinase gamma subunit (PRKAG3)
            mRNA, complete cds.
ACCESSION   AF214520
VERSION     AF214520.1  GI:8215683
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 1873)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A., Thelander,M.,
            Rogel-Gaillard,C., Paul,S., Iannuccelli,N., Rask,L., Ronne,H.,
            Lundstrom,K., Reinsch,N., Gellin,J., Kalm,E., Roy,P.L., Chardon,P.
            and Andersson,L.
  TITLE     A mutation in PRKAG3 associated with excess glycogen content in pig
            skeletal muscle
  JOURNAL   Science 288 (5469), 1248-1251 (2000)
  MEDLINE   20280150
   PUBMED   10818001
REFERENCE   2  (bases 1 to 1873)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A.,
            Rogel-Gaillard,C., Paul,S., Gellin,J., Lundstrom,K., Reinsch,N.,
            Kalm,E., Le Roy,P., Chardon,P. and Andersson,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-1999) Dept Animal Breeding and Genetics, Swedish
            University of Agricultural Sciences, BMC box 597, Uppsala 751 24,
            Sweden
FEATURES             Location/Qualifiers
     source          1. .1873
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /chromosome="15"
                     /map="15q"
                     /tissue_type="skeletal muscle"
     gene            1. .1873
                     /gene="PRKAG3"
     CDS             1. .1395
                     /gene="PRKAG3"
                     /note="AMPKG3"
                     /codon_start=1
                     /product="AMP-activated protein kinase gamma subunit"
                     /protein_id="AAF73988.1"
                     /db_xref="GI:8215684"
                     /translation="MSFLEQGESRSWPSRAVTTSSERSHGDQGNKASRWTRQEDVEEG
                     GPPGPREGPQSRPVAESTGQEATFPKATPLAQAAPLAEVDNPPTERDILPSDCAASAS
                     DSNTDHLDLGIEFSASAASGDELGLVEEKPAPCPSPEVLLPRLGWDDELQKPGAQVYM
                     HFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEEHKIETWREIYLQGCFKPLVSISPNDSLFEAVYALI
                     KNRIHRLPVLDPVSGAVLHILTHKRLLKFLHIFGTLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNETGQVVGLYSRFDVIHLAAQQTYNHLDMNV
                     GEALRQRTLCLEGVLSCQPHETLGEVIDRIVREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT      382 a    580 c    535 g    376 t
ORIGIN


  Query Match           71.7%;  Score 72.4;  DB 4;  Length 1873;
  Best Local Similarity  83.7%;  Pred. No. 8.2e-10;
  Matches   82;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      4 tccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcc 63
          |||||  |||  |  || |||| ||||||||||||||| || |||| ||| || || |||
Db    409 TCCCCAGAGGTGCTGTTACCCAGGCTGGGCTGGGATGATGAGCTGCAGAAGCCGGGGGCC 468

Qy     64 cagatctacatgcgcttcatgcaggagcacacctgcta 101
          ||| |||||||| |||||||||||||||||||||||||
Db    469 CAGGTCTACATGCACTTCATGCAGGAGCACACCTGCTA 506


RESULT  12
AX099800
LOCUS       AX099800               1873 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 27 from Patent WO0120003.
ACCESSION   AX099800
VERSION     AX099800.1  GI:13538834
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
```

```
REFERENCE   1  (bases 1 to 1873)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 27 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .1873
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
     CDS             1. .1395
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35800.1"
                     /db_xref="GI:13538835"
                     /translation="MSFLEQGESRSWPSRAVTTSSERSHGDQGNKASRWTRQEDVEEG
                     GPPGPREGPQSRPVAESTGQEATFPKATPLAQAAPLAEVDNPPTERDILPSDCAASAS
                     DSNTDHLDLGIEFSASAASGDELGLVEEKPAPCPSPEVLLPRLGWDDELQKPGAQVYM
                     HFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEEHKIETWREIYLQGCFKPLVSISPNDSLFEAVYALI
                     KNRIHRLPVLDPVSGAVLHILTHKRLLKFLHIFGTLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNETGQVVGLYSRFDVIHLAAQQTYNHLDMNV
                     GEALRQRTLCLEGVLSCQPHETLGEVIDRIVREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT      382 a    580 c    535 g    376 t
ORIGIN


  Query Match           71.7%;  Score 72.4;  DB 6;  Length 1873;
  Best Local Similarity  83.7%;  Pred. No. 8.2e-10;
  Matches   82;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       4 tccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaaccccggcgcc 63
           |||||  |||  |  || |||| |||||||||||||||| || |||| ||| || || |||
Db     409 TCCCCAGAGGTGCTGTTACCCAGGCTGGGCTGGGATGATGAGCTGCAGAAGCCGGGGGCC 468

Qy      64 cagatctacatgcgcttcatgcaggagcacacctgcta 101
           ||| ||||||||| ||||||||||||||||||||||||
Db     469 CAGGTCTACATGCACTTCATGCAGGAGCACACCTGCTA 506


RESULT  13
AX099804
LOCUS       AX099804               2022 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 31 from Patent WO0120003.
ACCESSION   AX099804
VERSION     AX099804.1  GI:13538838
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 2022)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 31 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .2022
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
BASE COUNT      412 a    623 c    593 g    394 t
ORIGIN


  Query Match           71.7%;  Score 72.4;  DB 6;  Length 2022;
  Best Local Similarity  83.7%;  Pred. No. 8.2e-10;
  Matches   82;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       4 tccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaaccccggcgcc 63
           |||||  |||  |  || |||| |||||||||||||||| || |||| ||| || || |||
Db     559 TCCCCAGAGGTGCTGTTACCCAGGCTGGGCTGGGATGATGAGCTGCAGAAGCCGGGGGCC 618

Qy      64 cagatctacatgcgcttcatgcaggagcacacctgcta 101
           ||| ||||||||| ||||||||||||||||||||||||
Db     619 CAGGTCTACATGCACTTCATGCAGGAGCACACCTGCTA 656


RESULT  14
AF214521
LOCUS       AF214521               5888 bp    DNA     linear   MAM 03-JUN-2000
DEFINITION  Sus scrofa AMPK gamma subunit (PRKAG3) gene, complete cds.
ACCESSION   AF214521
VERSION     AF214521.1  GI:8215685
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 5888)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A., Thelander,M.,
            Rogel-Gaillard,C., Paul,S., Iannuccelli,N., Rask,L., Ronne,H.,
            Lundstrom,K., Reinsch,N., Gellin,J., Kalm,E., Roy,P.L., Chardon,P.
            and Andersson,L.
  TITLE     A mutation in PRKAG3 associated with excess glycogen content in pig
            skeletal muscle
  JOURNAL   Science 288 (5469), 1248-1251 (2000)
  MEDLINE   20280150
   PUBMED   10818001
REFERENCE   2  (bases 1 to 5888)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A.,
            Rogel-Gaillard,C., Paul,S., Gellin,J., Lundstrom,K., Reinsch,N.,
            Kalm,E., Le Roy,P., Chardon,P. and Andersson,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-1999) Dept Animal Breeding and Genetics, Swedish
            University of Agricultural Sciences, BMC Box 597, Uppsala 751 24,
            Sweden
FEATURES             Location/Qualifiers
     source          1. .5888
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /chromosome="15"
                     /map="15q"
     mRNA            join(<1. .154,515. .918,1809. .1890,2349. .2407,2509. .2554,
                     2771. .2825,3027. .3153,3286. .3451,4578. .4615,4791. .4937,
                     5294. .>5410)
                     /gene="PRKAG3"
                     /product="AMPK gamma subunit"
     gene            <1. .>5410
                     /gene="PRKAG3"
     CDS             join(1. .154,515. .918,1809. .1890,2349. .2407,2509. .2554,
                     2771. .2825,3027. .3153,3286. .3451,4578. .4615,4791. .4937,
                     5294. .5410)
                     /gene="PRKAG3"
                     /codon_start=1
                     /product="AMPK gamma subunit"
                     /protein_id="AAF73989.1"
                     /db_xref="GI:8215686"
                     /translation="MSFLEQGESRSWPSRAVTTSSERSHGDQGNKASRWTRQEDVEEG
                     GPPGPREGPQSRPVAESTGQEATFPKATPLAQAAPLAEVDNPPTERDILPSDCAASAS
                     DSNTDHLDLGIEFSASAASGDELGLVEEKPAPCPSPEVLLPRLGWDDELQKPGAQVYM
                     HFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEEHKIETWREIYLQGCFKPLVSISPNDSLFEAVYALI
                     KNRIHRLPVLDPVSGAVLHILTHKRLLKFLHIFGTLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNETGQVVGLYSRFDVIHLAAQQTYNHLDMNV
                     GEALRQRTLCLEGVLSCQPHETLGEVIDRIVREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT     1274 a   1634 c   1638 g   1339 t      3 others
ORIGIN
```

```
  Query Match             71.7%;  Score 72.4;  DB 4;  Length 5888;
  Best Local Similarity   83.7%;  Pred. No. 7.6e-10;
  Matches   82;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       4 tccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcc 63
           |||||  |||  |  || |||| ||||||||||||||| || |||| ||| || || |||
Db     769 TCCCCAGAGGTGCTGTTACCCAGGCTGGGCTGGGATGATGAGCTGCAGAAGCCGGGGGCC 828

Qy      64 cagatctacatgcgcttcatgcaggagcacacctgcta 101
           ||| |||||||| |||||||||||||||||||||||||
Db     829 CAGGTCTACATGCACTTCATGCAGGAGCACACCTGCTA 866


RESULT  15
AF336381/c
LOCUS       AF336381              227724 bp    DNA     linear   HTG 02-APR-2001
DEFINITION  Mus musculus chromosome 1 clone PAC510; PAC457, *** SEQUENCING IN
            PROGRESS ***, 3 unordered pieces.
ACCESSION   AF336381
VERSION     AF336381.1  GI:13507298
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 227724)
  AUTHORS   Rump,A., Hayes,C., Brown,S.D.M. and Rosenthal,A.
  TITLE     Mouse chromosome 1 genomic sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 227724)
  AUTHORS   Rump,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JAN-2001) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenberstr. 11, Jena 07745, Germany
COMMENT     * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    17869: contig of 17869 bp in length
            *    17870    17969: gap of unknown length
            *    17970    32746: contig of 14777 bp in length
            *    32747    32846: gap of unknown length
            *    32847   227724: contig of 194878 bp in length.
FEATURES             Location/Qualifiers
     source          1..227724
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="1"
                     /clone="PAC510; PAC457"
BASE COUNT    57663 a  55423 c  56238 g  58199 t    201 others
ORIGIN


  Query Match             71.7%;  Score 72.4;  DB 2;  Length 227724;
  Best Local Similarity   83.7%;  Pred. No. 5.8e-10;
  Matches   82;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       4 tccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcc 63
           ||||| |||||||||| ||| |   ||| | |||||||||||||| | ||| ||||| |||
Db    8211 TCCCCACAGGCCCCACTTCTCGGCCTGAGTTGGGATGACGAACTTCAGAAGCCCGGAGCC 8152

Qy      64 cagatctacatgcgcttcatgcaggagcacacctgcta 101
           ||| |||||||||| |||||||||||| |||||||| ||
Db    8151 CAGGTCTACATGCACTTCATGCAGGAACACACCTGTTA 8114
```

```
Search completed: October  3, 2002, 14:49:54
Job time: 12114 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 16:30:43 ; Search time 719.93 Seconds
(without alignments)
240.868 Million cell updates/sec

Title: US-09-826-581-5_COPY_500_600
Perfect score: 101
Sequence: 1 ctgtccccgcaggccccatt.....atgcaggagcacacctgcta 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	101	100.0	989	22	AAH43682	PRKAG3 intron 2 -
2	101	100.0	1647	22	AAH43685	PRKAG3 cDNA. Homo
3	101	100.0	2109	22	AAD03296	Human AMPK gamma s
4	101	100.0	2115	22	AAD03320	Human AMPK gamma s
5	72.4	71.7	1867	22	AAD03295	Pig AMPK gamma sub
6	72.4	71.7	1873	22	AAD03319	Pig AMPK gamma sub
7	72.4	71.7	2022	22	AAD03321	Sus scrofa PRKAG3
8	30.2	29.9	1428	23	AAS51418	Pseudomonas aerugi
9	29.4	29.1	3918	23	AAS74198	DNA encoding novel
10	29.4	29.1	3918	23	AAS80424	DNA encoding novel
11	29.4	29.1	4397	23	AAS74446	DNA encoding novel
12	29.4	29.1	5122	23	AAS73340	DNA encoding novel
13	29	28.7	1406	23	ABL19505	Drosophila melanog
14	29	28.7	3406	23	ABL19504	Drosophila melanog
15	29	28.7	3593	23	ABL19502	Drosophila melanog
16	28.6	28.3	7816	22	AAK52934	Human polynucleoti
17	28.6	28.3	7818	22	AAK51950	Human polynucleoti
18	28.6	28.3	14533	23	AAS65883	DNA encoding novel
c 19	28	27.7	2175	22	AAL04858	Human reproductive
20	27.8	27.5	669	21	AAF12737	Aspergillus oryzae
c 21	27.6	27.3	2184	19	AAV34327	Human glycogenin-2
c 22	27.6	27.3	3184	19	AAV34328	Human glycogenin-2
c 23	27.6	27.3	3267	19	AAV34326	Human glycogenin-2
c 24	27.4	27.1	1319	23	ABL02979	Drosophila melanog
25	27.4	27.1	1803	23	ABL14873	Drosophila melanog
c 26	27.4	27.1	4514	23	ABL14872	Drosophila melanog
c 27	27.4	27.1	27671	21	AAA81474	N. meningitidis pa
c 28	27.4	27.1	349980	21	AAF21611	Neisseria meningit
c 29	27.2	26.9	876	22	AAK92035	Human cDNA 5'-end
c 30	27.2	26.9	876	22	AAK93756	Human cDNA clone r
c 31	27.2	26.9	1962	22	AAK94805	Human full-length
32	27	26.7	874	21	AAD02325	Human serine prote
33	27	26.7	1802	22	AAH89959	Human bone marrow
34	27	26.7	2365	20	AAZ34002	Human PRO351 nucle
35	27	26.7	2365	21	AAC78486	Human PRO351 (UNQ3
36	27	26.7	2365	21	AAC58230	Human PRO351 nucle
37	27	26.7	2664	18	AAT38449	cDNA for DNA polym
38	27	26.7	4466	23	ABL02978	Drosophila melanog
c 39	27	26.7	8899	22	ABA21236	Human nervous syst
40	26.8	26.5	2325	23	AAS54311	Pseudomonas aerugi
c 41	26.6	26.3	524	22	ABA61059	Human foetal liver
c 42	26.6	26.3	524	22	ABA28966	Probe #7432 for ge
c 43	26.6	26.3	524	22	AAK09352	Human brain expres
c 44	26.6	26.3	524	22	AAK35241	Human bone marrow
c 45	26.6	26.3	524	22	AAI40957	Probe #9643 used t

ALIGNMENTS

RESULT 1
AAH43682
ID AAH43682 standard; DNA; 989 BP.
XX
AC AAH43682;
XX
DT 21-JAN-2002 (first entry)
XX
DE PRKAG3 intron 2 - intron 4.
XX
KW Human; AMP-activated protein kinase gamma 3 subunit; PRKAG3; variant;
KW metabolic disease; diabetes; obesity; substitution; ds.
XX
OS Homo sapiens.
XX
FH Key. Location/Qualifiers
FT intron 1..21
FT /*tag= a
FT /number = "Intron 2"
FT /note= "3' portion of intron 2"
FT exon 22..177
FT /*tag= b
FT /number= "Exon 3"
FT intron 178..541
FT /*tag= c
FT /number= "Intron 3"
FT exon 542..945
FT /*tag= d
FT /number= "Exon 4"
FT intron 946..989
FT /*tag= e
FT /number= "Intron 4"

```
FT                   /note= "5' portion of intron 4"
XX
PN   WO200177305-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-SE00765.
XX
PR   07-APR-2000; 2000US-195665P.
XX
PA   (AREX-) AREXIS AB.
XX
PI   Andersson L,  Luthman H,  Marklund S;
XX
DR   WPI; 2001-657170/75.
XX
PT   New variants of human AMP-activated protein kinase gamma3 subunit
PT   associated with a metabolic disease e.g. diabetes or obesity and method
PT   for determining a risk estimate of diseases in subject by detecting the
PT   variant -
XX
PS   Example 1; Fig 2; 25pp; English.
XX
CC   The sequences given in AAH43681-84 represents genomic fragments
CC   encoding the human AMP-activated protein kinase gamma 3 subunit
CC   (PRKAG3). Detecting the presence of the PRKAG3 DNA, or a variant,
CC   is useful in determining a risk estimate of a metabolic disease,
CC   such as diabetes or obesity, in a subject.  The variation may occur
CC   in exons 3, 4 or 10.  In exon 3 variation may be a substitution of
CC   a G for a C at nucleotide 320, resulting in the amino acid
CC   substitution P71A; in exon 4 variation may be a substitution of a
CC   T for a C at nucleotide 550; and in exon 10 variation may be a
CC   substitution of a T for a C at nucleotide 1037, resulting in the
CC   amino acid substitution R340W.  There may also be nucleotide variation
CC   in intron 6.
XX
SQ   Sequence 989 BP; 229 A; 306 C; 286 G; 168 T; 0 other;


  Query Match             100.0%;  Score 101;  DB 22;  Length 989;
  Best Local Similarity   100.0%;  Pred. No. 9.8e-23;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     793 ctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggc 852

Qy      61 gcccagatctacatgcgcttcatgcaggagcacacctgcta 101
           |||||||||||||||||||||||||||||||||||||||||
Db     853 gcccagatctacatgcgcttcatgcaggagcacacctgcta 893


RESULT   2
AAH43685
ID   AAH43685 standard; cDNA; 1647 BP.
XX
AC   AAH43685;
XX
DT   21-JAN-2002  (first entry)
XX
DE   PRKAG3 cDNA.
XX
KW   Human; AMP-activated protein kinase gamma 3 subunit; PRKAG3; variant;
KW   metabolic disease; diabetes; obesity; substitution; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             20..1489
FT                   /*tag=  a
FT                   /product= "PRKAG3"
FT   variation       230
FT                   /*tag=  b
FT                   /label= "C230G"
FT                   /note= "Causes P71A"
FT   variation       559
FT                   /*tag=  c
FT                   /label= "T559C"
FT                   /note= "Silent variation"
FT   variation       1037
FT                   /*tag=  d
FT                   /label= "C1037T"
FT                   /note= "Causes R340W"
XX
PN   WO200177305-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-SE00765.
XX
PR   07-APR-2000; 2000US-195665P.
XX
PA   (AREX-) AREXIS AB.
XX
PI   Andersson L,  Luthman H,  Marklund S;
XX
DR   WPI; 2001-657170/75.
DR   P-PSDB; QQB47679.
XX
PT   New variants of human AMP-activated protein kinase gamma3 subunit
PT   associated with a metabolic disease e.g. diabetes or obesity and method
PT   for determining a risk estimate of diseases in subject by detecting the
PT   variant -
XX
PS   Disclosure; Fig 5; 25pp; English.
XX
CC   This sequence represents the full length cDNA encoding the human
CC   AMP-activated protein kinase gamma 3 subunit (PRKAG3). Detecting
CC   the presence of the PRKAG3 DNA, or a variant, is useful in determining
CC   a risk estimate of a metabolic disease, such as diabetes or obesity,
CC   in a subject.  The variation may occur in exons 3, 4 or 10.  In exon
CC   3 variation may be a substitution of a G for a C at nucleotide 320,
CC   resulting in the amino acid substitution P71A; in exon 4 variation may
CC   be a substitution of a T for a C at nucleotide 550; and in exon 10
CC   variation may be a substitution of a T for a C at nucleotide 1037,
CC   resulting in the amino acid substitution R340W.  There may also be
CC   nucleotide variation in intron 6.  The numbering of these
CC   variations is based on the full length cDNA as given, rather than on
CC   position 1 of the open reading frame.
XX
SQ   Sequence 1647 BP; 346 A; 502 C; 462 G; 337 T; 0 other;


  Query Match             100.0%;  Score 101;  DB 22;  Length 1647;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-22;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     500 ctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggc 559

Qy      61 gcccagatctacatgcgcttcatgcaggagcacacctgcta 101
           |||||||||||||||||||||||||||||||||||||||||
Db     560 gcccagatctacatgcgcttcatgcaggagcacacctgcta 600


RESULT   3
AAD03296
ID   AAD03296 standard; DNA; 2109 BP.
XX
AC   AAD03296;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Human AMPK gamma subunit muscle-specific isoform, PRKAG3 cDNA.
XX
```

```
KW   Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW   PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW   genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW   cystathione beta synthase; CBS; cardiant; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1..471
FT                   /*tag= a
FT   CDS             472..1389
FT                   /*tag= b
FT                   /product= "Human Prkag3 protein"
FT   3'UTR           1390..2109
FT                   /*tag= c
XX
PN   WO200120003-A2.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-EP09896.
XX
PR   10-SEP-1999;   99EP-0402236.
PR   18-MAY-2000; 2000EP-0401388.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA   (ANDE/) ANDERSSON L.
PA   (LOOF/) LOOFT C.
PA   (KALM/) KALM E.
XX
PI   Andersson L,  Looft C,  Kalm E,  Milan D,  Robic A,  Rogel-Gaillard C;
PI   Iannuccelli N,  Gellin J,  Le Roy P,  Chardon P;
XX
DR   WPI; 2001-244810/25.
DR   P-PSDB; AAE00221.
XX
PT   New variants of the gamma subunit of vertebrate adenosine
PT   monophosphate-activated kinase for diagnosis or treatment of disorders
PT   associated with energy metabolism such as diabetes, obesity, and
PT   myopathy  -
XX
PS   Claim 12; Fig 2; 71pp; English.
XX
CC   The present sequence is a cDNA encoding human adenosine monophosphate
CC   (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC   PRKAG3. Mutation in Prkag3 results in an altered regulation of
CC   carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is
CC   useful as therapeutic for treating carbohydrate metabolism disorders such
CC   as diabetes, obesity, and disorders associated with muscle metabolism
CC   such as myopathy and cardiovascular diseases, to modulate AMPK
CC   activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC   and its functionally altered mutants are useful for the diagnostic
CC   evaluation, genetic testing and prognosis of a metabolic disorder,
CC   preferably a carbohydrate metabolism disorder. Primers that can detect
CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
CC   from the expression of a functionally altered allele of PRKAG3.
CC   Transgenic animal and host cell transformed with PRKAG3 or a
CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC   screening compounds able to modulate AMPK activity. Nucleic acid
CC   encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
CC   of PRKAG3 and is useful in gene therapy.
XX
SQ   Sequence 2109 BP; 458 A; 621 C; 560 G; 470 T; 0 other;

  Query Match             100.0%;  Score 101;  DB 22;  Length 2109;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-22;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db     400 ctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggc 459

Qy      61 gcccagatctacatgcgcttcatgcaggagcacacctgcta 101
           |||||||||||||||||||||||||||||||||||||||||
Db     460 gcccagatctacatgcgcttcatgcaggagcacacctgcta 500


RESULT   4
AAD03320
ID   AAD03320 standard; cDNA; 2115 BP.
XX
AC   AAD03320;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Human AMPK gamma subunit muscle-specific isoform, complete PRKAG3 cDNA.
XX
KW   Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW   PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW   genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW   cystathione beta synthase; CBS; cardiant; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1395
FT                   /*tag= a
FT                   /product= "Human complete Prkag3 protein"
XX
PN   WO200120003-A2.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-EP09896.
XX
PR   10-SEP-1999;   99EP-0402236.
PR   18-MAY-2000; 2000EP-0401388.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA   (ANDE/) ANDERSSON L.
PA   (LOOF/) LOOFT C.
PA   (KALM/) KALM E.
XX
PI   Andersson L,  Looft C,  Kalm E,  Milan D,  Robic A,  Rogel-Gaillard C;
PI   Iannuccelli N,  Gellin J,  Le Roy P,  Chardon P;
XX
DR   WPI; 2001-244810/25.
DR   P-PSDB; AAE00223.
XX
PT   New variants of the gamma subunit of vertebrate adenosine
PT   monophosphate-activated kinase for diagnosis or treatment of disorders
PT   associated with energy metabolism such as diabetes, obesity, and
PT   myopathy  -
XX
PS   Claim 12; Page 65-68; 71pp; English.
XX
CC   The present sequence is a cDNA encoding human adenosine monophosphate
CC   (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC   complete PRKAG3. Mutation in Prkag3 results in an altered regulation of
CC   carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is
CC   useful as therapeutic for treating carbohydrate metabolism disorders such
CC   as diabetes, obesity, and disorders associated with muscle metabolism
CC   such as myopathy and cardiovascular diseases, to modulate AMPK
CC   activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC   and its functionally altered mutants are useful for the diagnostic
CC   evaluation, genetic testing and prognosis of a metabolic disorder,
CC   preferably a carbohydrate metabolism disorder. Primers that can detect
CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
CC   from the expression of a functionally altered allele of PRKAG3.
CC   Transgenic animal and host cell transformed with PRKAG3 or a
CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC   screening compounds able to modulate AMPK activity. Nucleic acid
```

```
CC   encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
CC   of PRKAG3 and is useful in gene therapy.
XX
SQ   Sequence 2115 BP; 460 A; 622 C; 562 G; 471 T; 0 other;


  Query Match             100.0%;  Score 101;  DB 22;  Length 2115;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-22;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 ctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     406 ctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggc 465

Qy      61 gcccagatctacatgcgcttcatgcaggagcacacctgcta 101
           |||||||||||||||||||||||||||||||||||||||||
Db     466 gcccagatctacatgcgcttcatgcaggagcacacctgcta 506


RESULT   5
AAD03295
ID   AAD03295 standard; cDNA; 1867 BP.
XX
AC   AAD03295;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Pig AMPK gamma subunit muscle-specific isoform, PRKAG3 cDNA.
XX
KW   Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW   PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW   genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW   cystathione beta synthase; CBS; cardiant; gene therapy; RN locus;
KW   chromosome 15; ss.
XX
OS   Sus scrofa.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1..471
FT                   /*tag= a
FT   CDS             472..1389
FT                   /*tag= b
FT                   /product= "Sus scrofa PRKAG3 protein"
FT   3'UTR           1390..1867
FT                   /*tag= c
XX
PN   WO200120003-A2.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-EP09896.
XX
PR   10-SEP-1999;   99EP-0402236.
PR   18-MAY-2000; 2000EP-0401388.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA   (ANDE/) ANDERSSON L.
PA   (LOOF/) LOOFT C.
PA   (KALM/) KALM E.
XX
PI   Andersson L,  Looft C,  Kalm E,  Milan D,  Robic A,  Rogel-Gaillard C;
PI   Iannuccelli N,  Gellin J,  Le Roy P,  Chardon P;
XX
DR   WPI; 2001-244810/25.
DR   P-PSDB; AAE00220.
XX
PT   New variants of the gamma subunit of vertebrate adenosine
PT   monophosphate-activated kinase for diagnosis or treatment of disorders
PT   associated with energy metabolism such as diabetes, obesity, and
PT   myopathy -
XX
PS   Claim 12; Fig 2; 71pp; English.
XX
CC   The present sequence is a cDNA encoding pig adenosine monophosphate
CC   (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC   PRKAG3. Prkag3 gene is located in the RN locus of chromosome 15.
CC   Mutation in Prkag3 results in an altered regulation of carbohydrate
CC   metabolism, particularly in skeletal muscle. PRKAG3 is useful as
CC   therapeutic for treating carbohydrate metabolism disorders such as
CC   diabetes, obesity, and disorders associated with muscle metabolism
CC   such as myopathy and cardiovascular diseases, to modulate AMPK
CC   activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC   and its functionally altered mutants are useful for the diagnostic
CC   evaluation, genetic testing and prognosis of a metabolic disorder,
CC   preferably a carbohydrate metabolism disorder. Primers that can detect
CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
CC   from the expression of a functionally altered allele of PRKAG3.
CC   Transgenic animal and host cell transformed with PRKAG3 or a
CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC   screening compounds able to modulate AMPK activity. Nucleic acid
CC   encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
CC   of PRKAG3 and is useful in gene therapy.
XX
SQ   Sequence 1867 BP; 380 A; 583 C; 529 G; 375 T; 0 other;


  Query Match             71.7%;  Score 72.4;  DB 22;  Length 1867;
  Best Local Similarity   83.7%;  Pred. No. 1.3e-13;
  Matches   82;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy       4 tccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcc 63
           |||||  |||  |  || |||| ||||||||||||||| || |||| ||| || || |||
Db     403 tccccagaggtgctgttacccaggctgggctgggatgatgagctgcagaagccgggggcc 462

Qy      64 cagatctacatgcgcttcatgcaggagcacacctgcta 101
           ||| |||||||||| ||||||||||||||||||||||||
Db     463 caggtctacatgcacttcatgcaggagcacacctgcta 500


RESULT   6
AAD03319
ID   AAD03319 standard; cDNA; 1873 BP.
XX
AC   AAD03319;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Pig AMPK gamma subunit muscle-specific isoform, complete PRKAG3 cDNA.
XX
KW   Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW   PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW   genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW   cystathione beta synthase; CBS; cardiant; gene therapy; RN locus;
KW   chromosome 15; ss.
XX
OS   Sus scrofa.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1395
FT                   /*tag= a
FT                   /product= "Sus scrofa complete Prkag3 protein"
XX
PN   WO200120003-A2.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-EP09896.
XX
PR   10-SEP-1999;   99EP-0402236.
PR   18-MAY-2000; 2000EP-0401388.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA   (ANDE/) ANDERSSON L.
```

PA (LOOF/) LOOFT C.
PA (KALM/) KALM E.
XX
PI Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;
PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;
XX
DR WPI; 2001-244810/25.
DR P-PSDB; AAE00222.
XX
PT New variants of the gamma subunit of vertebrate adenosine
PT monophosphate-activated kinase for diagnosis or treatment of disorders
PT associated with energy metabolism such as diabetes, obesity, and
PT myopathy -
XX
PS Claim 12; Page 62-64; 71pp; English.
XX
CC The present sequence is a cDNA encoding pig adenosine monophosphate
CC (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC complete PRKAG3. Prkag3 gene is located in the RN locus of chromosome
CC 15. Mutation in Prkag3 results in an altered regulation of carbohydrate
CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as
CC therapeutic for treating carbohydrate metabolism disorders such as
CC diabetes, obesity, and disorders associated with muscle metabolism
CC such as myopathy and cardiovascular diseases, to modulate AMPK
CC activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC and its functionally altered mutants are useful for the diagnostic
CC evaluation, genetic testing and prognosis of a metabolic disorder,
CC preferably a carbohydrate metabolism disorder. Primers that can detect
CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC useful for detecting a dysfunction of carbohydrate metabolism resulting
CC from the expression of a functionally altered allele of PRKAG3.
CC Transgenic animal and host cell transformed with PRKAG3 or a
CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC screening compounds able to modulate AMPK activity. Nucleic acid
CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC in a sequence encoding the first cystathione beta synthase (CBS) domain
CC of PRKAG3 and is useful in gene therapy.
XX
SQ Sequence 1873 BP; 382 A; 580 C; 535 G; 376 T; 0 other;

Query Match 71.7%; Score 72.4; DB 22; Length 1873;
Best Local Similarity 83.7%; Pred. No. 1.3e-13;
Matches 82; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy      4 tccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcc 63
          |||||  |||  |  || |||| ||||||||||||||| || |||| ||| || || |||
Db    409 tccccagaggtgctgttacccaggctgggctgggatgatgagctgcagaagccgggggcc 468

Qy     64 cagatctacatgcgcttcatgcaggagcacacctgcta 101
          ||| ||||||||| ||||||||||||||||||||||||
Db    469 caggtctacatgcacttcatgcaggagcacacctgcta 506
```

RESULT 7
AAD03321
ID AAD03321 standard; DNA; 2022 BP.
XX
AC AAD03321;
XX
DT 13-JUN-2001 (first entry)
XX
DE Sus scrofa PRKAG3 splice variant DNA.
XX
KW Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW cystathione beta synthase; CBS; cardiant; gene therapy; ds.
XX
OS Sus scrofa.
XX
FH Key Location/Qualifiers
FT CDS 1..1545
FT /*tag= a
FT /product= "Sus scrofa Prkag3 splice variant"
XX
PN WO200120003-A2.
XX
PD 22-MAR-2001.
XX
PF 11-SEP-2000; 2000WO-EP09896.
XX
PR 10-SEP-1999; 99EP-0402236.
PR 18-MAY-2000; 2000EP-0401388.
XX
PA (INRG) INRA INST NAT RECH AGRONOMIQUE.
PA (ANDE/) ANDERSSON L.
PA (LOOF/) LOOFT C.
PA (KALM/) KALM E.
XX
PI Andersson L, Looft C, Kalm E, Milan D, Robic A, Rogel-Gaillard C;
PI Iannuccelli N, Gellin J, Le Roy P, Chardon P;
XX
DR WPI; 2001-244810/25.
DR P-PSDB; AAE00224.
XX
PT New variants of the gamma subunit of vertebrate adenosine
PT monophosphate-activated kinase for diagnosis or treatment of disorders
PT associated with energy metabolism such as diabetes, obesity, and
PT myopathy -
XX
PS Claim 12; Page 69; 71pp; English.
XX
CC The present sequence is pig adenosine monophosphate (AMP)-activated
CC kinase (AMPK) gamma subunit muscle-specific isoform, PRKAG3 splice
CC variant DNA. Prkag3 gene is located in the RN locus of chromosome 15.
CC Mutation in Prkag3 results in an altered regulation of carbohydrate
CC metabolism, particularly in skeletal muscle. PRKAG3 is useful as
CC therapeutic for treating carbohydrate metabolism disorders such as
CC diabetes, obesity, and disorders associated with muscle metabolism
CC such as myopathy and cardiovascular diseases, to modulate AMPK
CC activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC and its functionally altered mutants are useful for the diagnostic
CC evaluation, genetic testing and prognosis of a metabolic disorder,
CC preferably a carbohydrate metabolism disorder. Primers that can detect
CC a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC useful for detecting a dysfunction of carbohydrate metabolism resulting
CC from the expression of a functionally altered allele of PRKAG3.
CC Transgenic animal and host cell transformed with PRKAG3 or a
CC heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC screening compounds able to modulate AMPK activity. Nucleic acid
CC encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC in a sequence encoding the first cystathione beta synthase (CBS) domain
CC of PRKAG3 and is useful in gene therapy.
XX
SQ Sequence 2022 BP; 412 A; 623 C; 593 G; 394 T; 0 other;

Query Match 71.7%; Score 72.4; DB 22; Length 2022;
Best Local Similarity 83.7%; Pred. No. 1.3e-13;
Matches 82; Conservative 0; Mismatches 16; Indels 0; Gaps 0;

```
Qy      4 tccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcc 63
          |||||  |||  |  || |||| ||||||||||||||| || |||| ||| || || |||
Db    559 tccccagaggtgctgttacccaggctgggctgggatgatgagctgcagaagccgggggcc 618

Qy     64 cagatctacatgcgcttcatgcaggagcacacctgcta 101
          ||| ||||||||| ||||||||||||||||||||||||
Db    619 caggtctacatgcacttcatgcaggagcacacctgcta 656
```

RESULT 8
AAS51418
ID AAS51418 standard; DNA; 1428 BP.
XX
AC AAS51418;

```
XX
DT   13-FEB-2002  (first entry)
XX
DE   Pseudomonas aeruginosa DNA for cellular proliferation protein #3.
XX
KW   Antisense; ds; prokaryotic cellular proliferation gene;
KW   antibiotic; antibacterial; drug design.
XX
OS   Pseudomonas aeruginosa.
XX
PN   WO200170955-A2.
XX
PD   27-SEP-2001.
XX
PF   21-MAR-2001; 2001WO-US09180.
XX
PR   21-MAR-2000; 2000US-191078P.
PR   23-MAY-2000; 2000US-206848P.
PR   26-MAY-2000; 2000US-207727P.
PR   23-OCT-2000; 2000US-242578P.
PR   27-NOV-2000; 2000US-253625P.
PR   22-DEC-2000; 2000US-257931P.
PR   16-FEB-2001; 2001US-269308P.
XX
PA   (ELIT-) ELITRA PHARM INC.
XX
PI   Haselbeck R,  Ohlsen KL,  Zyskind JW,  Wall D,  Trawick JD,  Carr GJ;
PI   Yamamoto RT,  Xu HH;
XX
DR   WPI; 2001-611495/70.
DR   P-PSDB; AAU33559.
XX
PT   New polynucleotides for the identification and development of
PT   antibiotics, comprise sequences of antisense nucleic acids -
XX
PS   Claim 27; Seq ID No 4000; 511pp; English.
XX
CC   The invention relates to antisense inhibitors of genes essential to
CC   prokaryotic cellular proliferation, their use in identifying the
CC   genes, their use in the discovery of novel antibiotics, the essential
CC   genes themselves and the encoded proteins. The prokaryotes used are
CC   Escherichia coli, Staphylococcus aureus, Salmonella typhi, Klebsiella
CC   pneumoniae, Pseudomonas aeruginosa and Enterococcus faecalis. The
CC   invention is also useful for the identification of potential new targets
CC   for antibiotic development. The antisense nucleic acids can also be used
CC   to identify proteins used in proliferation, to express these proteins,
CC   and to obtain antibodies capable of binding to the expressed proteins.
CC   The proteins can be used to screen compounds in rational drug discovery
CC   programmes. The antisense nucleic acid sequence is also useful to screen
CC   for homologous nucleic acids which are required for cell proliferation in
CC   a wide variety of organisms. The present sequence encodes an
CC   essential prokaryotic cellular proliferation protein.
CC   Note: The sequence data for this patent did not form part
CC   of the printed specification, but was obtained in electronic
CC   format directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1428 BP; 206 A; 525 C; 423 G; 274 T; 0 other;


  Query Match             29.9%;  Score 30.2;  DB 23;  Length 1428;
  Best Local Similarity   58.2%;  Pred. No. 3.2;
  Matches   53;  Conservative    0;  Mismatches   38;  Indels    0;  Gaps    0;

Qy       3 gtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgc 62
          || ||    |   | || |||| |||| ||||| |     ||| |    || |||
Db    1218 gttccgtatgtggctatatccctggctgacctgggcggtgatcctgttcatcgtcgccgc 1277

Qy      63 ccagatctacatgcgcttcatgcaggagcac 93
           | || |  |||||   |||||| |||||||
Db    1278 gctgagcatcatgctgatcatgccggagcac 1308
```

```
RESULT   9
AAS74198
ID   AAS74198 standard; cDNA; 3918 BP.
XX
AC   AAS74198;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #10002.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG10011.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 10002; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 3918 BP; 1328 A; 872 C; 822 G; 896 T; 0 other;


  Query Match             29.1%;  Score 29.4;  DB 23;  Length 3918;
  Best Local Similarity   56.8%;  Pred. No. 6.9;
  Matches   54;  Conservative    0;  Mismatches   41;  Indels    0;  Gaps    0;

Qy       4 tccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcc 63
           |||||  |  ||||||   ||  |||| | || |  | || || |||     | ||||
Db    2710 tcccctttatcccccatcctccctgctgaggtgtgtgggcgcacagcgcgggactggcggg 2769

Qy      64 cagatctacatgcgcttcatgcaggagcacacctg 98
           ||| || || |||    |  |   |||  |  |||
```

```
Db   2770 cagctccacctgcagccccagtgcgagatccactg 2804


RESULT  10
AAS80424
ID   AAS80424 standard; cDNA; 3918 BP.
XX
AC   AAS80424;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #16228.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG16237.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 16228; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 3918 BP; 1328 A; 872 C; 822 G; 896 T; 0 other;


  Query Match             29.1%;  Score 29.4;  DB 23;  Length 3918;
  Best Local Similarity   56.8%;  Pred. No. 6.9;
  Matches   54;  Conservative    0;  Mismatches   41;  Indels    0;  Gaps    0;

Qy       4 tccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcc 63
           |||||  |  ||||||   ||  |||| | || |  | || || |||     | ||||
Db    2710 tccccttatcccccatcctccctgctgaggtgtgtgggcgcacagcgcgggactggcggg 2769

Qy      64 cagatctacatgcgcttcatgcaggagcacacctg 98
           ||| || || |||     |  |   |||  |  |||
Db    2770 cagctccacctgcagccccagtgcgagatccactg 2804


RESULT  11
AAS74446
ID   AAS74446 standard; cDNA; 4397 BP.
XX
AC   AAS74446;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #10250.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT,  Liu C,  Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG10259.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity  -
XX
PS   Claim 1; SEQ ID No 10250; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 4397 BP; 1506 A; 980 C; 935 G; 976 T; 0 other;


  Query Match             29.1%;  Score 29.4;  DB 23;  Length 4397;
  Best Local Similarity   56.8%;  Pred. No. 7;
  Matches   54;  Conservative    0;  Mismatches   41;  Indels    0;  Gaps    0;
```

```
Qy       4 tccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcc 63
           |||||  |  ||||||   ||  |||| | || |  | || || |||     | ||||
Db    3189 tccccttatcccccatcctccctgctgaggtgtgtgggcgcacagcgcgggactggcggg 3248

Qy      64 cagatctacatgcgcttcatgcaggagcacacctg 98
           ||| || || |||    |  |  |||  |  |||
Db    3249 cagctccacctgcagccccagtgcgagatccactg 3283
```

RESULT 12
AAS73340
ID AAS73340 standard; cDNA; 5122 BP.
XX
AC AAS73340;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #9144.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR P-PSDB; ABG09153.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 1; SEQ ID No 9144; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 5122 BP; 1815 A; 1151 C; 1032 G; 1124 T; 0 other;

Query Match 29.1%; Score 29.4; DB 23; Length 5122;
Best Local Similarity 56.8%; Pred. No. 7.2;
Matches 54; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

```
Qy       4 tccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcc 63
           |||||  |  ||||||   ||  |||| | || |  | || || |||     | ||||
Db    3914 tccccttatcccccatcctccctgctgaggtgtgtgggcgcacagcgcgggactggcggg 3973

Qy      64 cagatctacatgcgcttcatgcaggagcacacctg 98
           ||| || || |||    |  |  |||  |  |||
Db    3974 cagctccacctgcagccccagtgcgagatccactg 4008
```

RESULT 13
ABL19505
ID ABL19505 standard; DNA; 1406 BP.
XX
AC ABL19505;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 9988.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 9988; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1406 BP; 358 A; 409 C; 404 G; 235 T; 0 other;

Query Match 28.7%; Score 29; DB 23; Length 1406;
Best Local Similarity 71.7%; Pred. No. 7.6;
Matches 38; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```
Qy      49 cggaaacccggcgcccagatctacatgcgcttcatgcaggagcacacctgcta 101
           | ||| ||||||||||| | | | |  ||||| | | |||||||| | ||  |||
Db     264 ctgaagcccggcgcccggctgtgcccgcgctgctttcaggagctctccgacta 316
```

RESULT 14

```
ABL19504
ID   ABL19504 standard; DNA; 3406 BP.
XX
AC   ABL19504;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster genomic polynucleotide SEQ ID NO 9985.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ds.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Claim 1; SEQ ID NO 9985; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 3406 BP; 902 A; 852 C; 828 G; 824 T; 0 other;


  Query Match             28.7%;  Score 29;  DB 23;  Length 3406;
  Best Local Similarity   71.7%;  Pred. No. 9;
  Matches   38;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      49 cggaaacccggcgcccagatctacatgcgcttcatgcaggagcacacctgcta 101
           | ||| |||||||||| | | | |  ||||| | | ||||||| | ||  |||
Db    1264 ctgaagcccggcgcccggctgtgcccgcgctgctttcaggagctctccgacta 1316


RESULT  15
ABL19502
ID   ABL19502 standard; DNA; 3593 BP.
XX
AC   ABL19502;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster genomic polynucleotide SEQ ID NO 9979.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ds.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Claim 1; SEQ ID NO 9979; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 3593 BP; 920 A; 953 C; 902 G; 818 T; 0 other;


  Query Match             28.7%;  Score 29;  DB 23;  Length 3593;
  Best Local Similarity   71.7%;  Pred. No. 9.1;
  Matches   38;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      49 cggaaacccggcgcccagatctacatgcgcttcatgcaggagcacacctgcta 101
           | ||| |||||||||| | | | |  ||||| | | ||||||| | ||  |||
Db    3377 ctgaagcccggcgcccggctgtgcccgcgctgctttcaggagctctccgacta 3429


Search completed: October  3, 2002, 16:30:46
Job time: 14316 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 16:15:04 ; Search time 5701.1 Seconds
(without alignments)
239.110 Million cell updates/sec

Title: US-09-826-581-5_COPY_500_600
Perfect score: 101
Sequence: 1 ctgtccccgcaggccccatt.....atgcaggagcacacctgcta 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	72.2	71.5	444	10	BF890374	BF890374 291826 MA
	2	72.2	71.5	548	10	BI775360	BI775360 467815 MA
	3	67.6	66.9	685	9	BB630381	BB630381 BB630381
c	4	31.4	31.1	326	9	AI106144	AI106144 db03d12_p
c	5	30.6	30.3	679	10	BE540087	BE540087 601060879
c	6	30.4	30.1	699	10	BI199916	BI199916 602760950
c	7	30.2	29.9	580	12	BH140228	BH140228 ZMMBBb000
c	8	30.2	29.9	894	12	CNS01MA1	AL150666 Anopheles
	9	29.8	29.5	910	12	CNS02SYT	AL212510 Tetraodon
	10	29.8	29.5	1010	12	CNS038S6	AL233007 Tetraodon
	11	29.6	29.3	655	10	BM133818	BM133818 NXLV_012_
	12	29.6	29.3	930	10	BE269692	BE269692 601185836
	13	29.4	29.1	955	10	BG394974	BG394974 602457369
	14	29.2	28.9	1000	10	BM468546	BM468546 AGENCOURT
	15	29	28.7	313	9	AA264931	AA264931 LD08544.5
c	16	29	28.7	436	10	T95291	T95291 ye43b03.s1
	17	29	28.7	546	9	AA246471	AA246471 LD05631.5
	18	29	28.7	554	9	AI260567	AI260567 LP04480.5
	19	29	28.7	561	10	BI229683	BI229683 RE27517.5
	20	29	28.7	583	10	BI232284	BI232284 RE28330.5
	21	29	28.7	589	9	AA539970	AA539970 LD19281.5
	22	29	28.7	627	10	BI633071	BI633071 SD27042.5
	23	29	28.7	629	9	AI259247	AI259247 LP02620.5
	24	29	28.7	666	10	BI242279	BI242279 RE39577.5
	25	29	28.7	738	9	AI544103	AI544103 LD46336.5
c	26	28.8	28.5	569	10	BG414376	BG414376 HVSMEk000
c	27	28.8	28.5	750	10	BM019844	BM019844 603648189
	28	28.6	28.3	339	10	BF773519	BF773519 CM2-IT003
	29	28.6	28.3	445	10	BG558441	BG558441 RH1Z2_67_
c	30	28.6	28.3	448	9	AW497449	AW497449 660052B03
	31	28.6	28.3	509	10	BG560394	BG560394 RH1Z2_73_
	32	28.6	28.3	591	9	AW678584	AW678584 WS1_16_E0
c	33	28.6	28.3	593	12	BH140663	BH140663 ZMMBBb000
	34	28.6	28.3	605	9	AA195150	AA195150 zr34c12.r
c	35	28.6	28.3	707	12	AG130442	AG130442 Pan trogl
	36	28.6	28.3	718	9	AV753229	AV753229 AV753229
	37	28.6	28.3	742	10	BE394076	BE394076 601312450
	38	28.6	28.3	904	10	BF033403	BF033403 601457923
c	39	28.4	28.1	260	9	AW816448	AW816448 QV4-ST023
c	40	28.4	28.1	712	9	BB623242	BB623242 BB623242
c	41	28.4	28.1	812	9	AA519383	AA519383 TgESTzz54
	42	28.2	27.9	517	9	AW631272	AW631272 hh80f03.y
c	43	28.2	27.9	525	9	AA922163	AA922163 om42b02.s
	44	28.2	27.9	592	12	AZ870322	AZ870322 2M0182F14
	45	28.2	27.9	802	12	BH077939	BH077939 RPCI-24-2

ALIGNMENTS

RESULT 1
BF890374
LOCUS BF890374 444 bp mRNA linear EST 25-APR-2001
DEFINITION 291826 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BF890374
VERSION BF890374.1 GI:12281760
KEYWORDS EST.
SOURCE cow.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea; Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 444)
AUTHORS Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keele,J.W.
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
JOURNAL Genome Res. 11 (4), 626-630 (2001)
MEDLINE 21180013
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR PRimers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCCAGTCACGACG
Plate: 57 row: H column: 11
Seq primer: ATTTAGGTGACACTATAG.
FEATURES Location/Qualifiers
source 1. .444
/organism="Bos taurus"
/db_xref="taxon:9913"

```
                     /clone_lib="MARC 3BOV"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled tissue from marrow, alveolar
                     macrophage, ovary, fetal semitendonosus muscle, and fetal
                     longissimus muscle."
BASE COUNT       91 a    147 c    139 g     67 t
ORIGIN


  Query Match             71.5%;  Score 72.2;  DB 10;  Length 444;
  Best Local Similarity   82.2%;  Pred. No. 7.5e-11;
  Matches   83;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 ctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggc 60
           | ||||||||||  ||  | | || ||| |||||||| || || |||||||| || ||
Db     295 CCGTCCCCGCAGGTGCCGGTGCTCAGGCTCGGCTGGGACGATGAGCTGCGGAAGCCAGGG 354

Qy      61 gcccagatctacatgcgcttcatgcaggagcacacctgcta 101
           || ||| ||||||||| |||||||||||||||||||||||
Db     355 GCTCAGGTCTACATGCACTTCATGCAGGAGCACACCTGCTA 395


RESULT   2
BI775360
LOCUS       BI775360                548 bp    mRNA    linear   EST 25-SEP-2001
DEFINITION  467815 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION   BI775360
VERSION     BI775360.1  GI:15776346
KEYWORDS    EST.
SOURCE      cow.
  ORGANISM  Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Bovinae; Bos.
REFERENCE   1  (bases 1 to 548)
  AUTHORS   Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
            Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett
            ,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G.,
            Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and
            Keele,J.W.
  TITLE     Sequence evaluation of four pooled-tissue normalized bovine cDNA
            libraries and construction of a gene index for cattle
  JOURNAL   Genome Res. 11 (4), 626-630 (2001)
  MEDLINE   21180013
COMMENT     Contact: Smith TPL
            USDA, ARS, US Meat Animal Research Center
            PO Box 166, Clay Center, NE 68933-0166, USA
            Tel: 402 762 4366
            Fax: 402 762 4390
            Email: smith@email.marc.usda.gov
            Single pass sequencing. Bases called and alt_trimmed with phred
            v0.980904.e. Vector identified by cross_match with the -minscore 18
            and -minmatch 12 options.
            PCR PRimers
            FORWARD: AGGAAACAGCTATGACCAT
            BACKWARD: GTTTTCCCAGTCACGACG
            Plate: 90  row: I  column: 16
            Seq primer: ATTTAGGTGACACTATAG.
FEATURES             Location/Qualifiers
     source          1. .548
                     /organism="Bos taurus"
                     /db_xref="taxon:9913"
                     /clone_lib="MARC 2BOV"
                     /tissue_type="pooled"
                     /lab_host="DH10B"
                     /note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;
                     Library made from pooled tissue from testis, thymus,
                     semitendonosus muscle, longissimus muscle, pancreas,
                     adrenal, and endometrium."
BASE COUNT      119 a    173 c    163 g     93 t
ORIGIN


  Query Match             71.5%;  Score 72.2;  DB 10;  Length 548;
  Best Local Similarity   82.2%;  Pred. No. 7.9e-11;
  Matches   83;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy       1 ctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggc 60
           | ||||||||||  ||  | | || ||| |||||||| || || |||||||| || ||
Db     337 CCGTCCCCGCAGGTGCCGGTGCTCATGCTCGGCTGGGACGATGAGCTGCGGAAACCATGG 396

Qy      61 gcccagatctacatgcgcttcatgcaggagcacacctgcta 101
           || ||| ||||||||| |||||||||||||||||||||||
Db     397 GCTCAGGTCTACATGCACTTCATGCAGGAGCACACCTGCTA 437


RESULT   3
BB630381
LOCUS       BB630381                685 bp    mRNA    linear   EST 26-OCT-2001
DEFINITION  BB630381 RIKEN full-length enriched, 6 days neonate skin Mus
            musculus cDNA clone A030014A04 5', mRNA sequence.
ACCESSION   BB630381
VERSION     BB630381.1  GI:16467364
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 685)
  AUTHORS   Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.
            , Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
            ,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
            Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
            ,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
            Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
            Muramatsu,M. and Hayashizaki,Y.
  TITLE     RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Yoshihide Hayashizaki
            Laboratory for Genome Exploration Research Group, RIKEN Genomic
            Sciences Center(GSC), Yokohama Institute
            The Institute of Physical and Chemical Research (RIKEN)
            1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
            Tel: 81-45-503-9222
            Fax: 81-45-503-9216
            Email: genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/
            Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
            ,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
             Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new
            genes. Genome Res. . 10 (10), 1617-1630 (2000)
             wagi,K., Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
            Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
            ,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
            Hayashizaki,Y.
             RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer. Genome Res. .
            10 (11), 1757-1771 (2000)
             Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
            ,Y. and Hayashizaki,Y.
             Computer-based methods for the mouse full-length cDNA
            encyclopedia: real-time sequence clustering for construction of a
            nonredundant cDNA library. Genome Res. . 11 (2), 281-289 (2001)
             Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa
            ,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
            Hayashizaki,Y.
             Computational Analysis of Full-Length Mouse cDNAs Compared with
            Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
             Please visit our web site (http://genome.gsc.riken.go.jp) for
            further details.
             e mouse tissues.
FEATURES             Location/Qualifiers
     source          1. .685
```

```
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="A030014A04"
                     /clone_lib="RIKEN full-length enriched, 6 days neonate
                     skin"
                     /tissue_type="skin"
                     /dev_stage="6 days neonate"
                     /lab_host="DH10B"
                     /note="Site_1: SalI; Site_2: BamHI; cDNA library was
                     prepared and sequenced in Mouse Genome Encyclopedia
                     Project of Genome Exploration Research Group in Riken
                     Genomic Sciences Center and Genome Science Laboratory in
                     RIKEN. Division of Experimental Animal Research in Riken
                     contributed to prepare mouse tissues. 1st strand cDNA was
                     primed with a primer [5'
                     GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
                     prepared by using trehalose thermo-activated reverse
                     transcriptase and subsequently enriched for full-length by
                     cap-trapper. cDNA went through one round of normalization
                     to Rot = 10.0 and subtraction to Rot = 185.0. Second
                     strand cDNA was prepared with the primer adapter of
                     sequence [5' GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCCC
                     3']. cDNA was cleaved with XhoI and BamHI. Vector: a
                     modified pBluescript KS(+) after bulk excision from Lambda
                     FLC I. "
BASE COUNT       165 a    199 c    198 g    123 t
ORIGIN


  Query Match             66.9%;  Score 67.6;  DB 9;  Length 685;
  Best Local Similarity   80.6%;  Pred. No. 1.7e-09;
  Matches   79;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy       4 tccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcc 63
           ||||| |||||||||| |||     ||| | |||||||||| |||| |  |||||||| |||
Db     496 TCCCCACAGGCCCCACTTCTTGGCCTGAGTTGGGATGACCAACTTCACAAACCCGGAGCC 555

Qy      64 cagatctacatgcgcttcatgcaggagcacacctgcta 101
           ||| |||||||||| ||| |||||||| |||||||| ||
Db     556 CAGGTCTACATGCACTTTATGCAGGAACACACCTGTTA 593


RESULT   4
AI106144/c
LOCUS       AI106144                 326 bp    mRNA    linear   EST 25-AUG-1998
DEFINITION  db03d12_p1 ZF adult heart library Danio rerio cDNA 5 prime similar
            to PYRUVATE KINASE, MUSCLE ISOZYME, mRNA sequence.
ACCESSION   AI106144
VERSION     AI106144.1  GI:3461247
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 326)
  AUTHORS   Chen,J.N., DeSauvage,F., Hosobuchi,M., Jackson,D.G. and Fishman
            ,M.C.
  TITLE     Expressed Sequences from The Adult Zebrafish Heart
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Mark C. Fishman
            Cardiovascular Research Center
            Massachusetts General Hospital
            Mail code 1494100A, 149 13th Street, Charlestown, MA 02129, USA
            Fax: 6177265806
            Email: fishman@mgh.cvrc.harvard.edu
            http://zebrafish.mgh.harvard.edu
            The original clones used for sequencing are no longer available;
            the library is available from Mark C. Fishman.
            Insert Length: 326   Std Error: 0.00
            Seq primer: p1.
FEATURES             Location/Qualifiers
     source          1. .326
                     /organism="Danio rerio"
                     /strain="AB"
                     /db_xref="taxon:7955"
                     /clone_lib="ZF adult heart library"
                     /sex="mixed"
                     /tissue_type="myocardium, endocardium, vessel"
                     /dev_stage="adult"
                     /lab_host="E. coli XL1 Blue"
                     /note="Organ: heart; Vector: LambdaZAPII; Site_1: EcoRI;
                     Site_2: XhoI"
BASE COUNT       83 a     66 c     96 g     79 t      2 others
ORIGIN


  Query Match             31.1%;  Score 31.4;  DB 9;  Length 326;
  Best Local Similarity   62.7%;  Pred. No. 34;
  Matches   47;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      23 ccaagctgggctgggatgacgaactgcggaaacccggcgcccagatctacatgcgcttca 82
           |||| |||          ||| |  || |  | || | |||  ||| || ||||  ||| ||
Db     256 CCAAACTGAAGGTCCTTGATGTCCTNCNCAGACACTGCGGGCAGGTCAACATTGGCTCCA 197

Qy      83 tgcaggagcacacct 97
           |||||  |||||||
Db     196 GGCAGGTTCACACCT 182


RESULT   5
BE540087/c
LOCUS       BE540087                 679 bp    mRNA    linear   EST 09-AUG-2000
DEFINITION  601060879F2 NIH_MGC_10 Homo sapiens cDNA clone IMAGE:3447378 5',
            mRNA sequence.
ACCESSION   BE540087
VERSION     BE540087.1  GI:9768732
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 679)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: Incyte Genomics, Inc.
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM8420  row: n  column: 19
            High quality sequence stop: 668.
FEATURES             Location/Qualifiers
     source          1. .679
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3447378"
                     /clone_lib="NIH_MGC_10"
                     /cell_line="MGC36"
                     /lab_host="DH10B"
                     /note="Organ: cervix; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.5 kb. Library prepared by Life
                     Technologies."
BASE COUNT      124 a    226 c    177 g    152 t
ORIGIN


  Query Match             30.3%;  Score 30.6;  DB 10;  Length 679;
  Best Local Similarity   62.3%;  Pred. No. 68;
  Matches   48;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;
```

```
Qy       9 gcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcccagat 68
           ||| |||   ||||   || || || || |   ||||||| ||| | | || |||  |||
Db     244 GCAAGCCGGCTTTCTGGAGGTGTGCAGGAAGAGCGAACTGGGGACAGCTGGGGCCTGGAT 185

Qy      69 ctacatgcgcttcatgc 85
             ||  |   | |||||
Db     184 TCACTGGACATACATGC 168


RESULT   6
BI199916/c
LOCUS       BI199916                 699 bp    mRNA    linear   EST 10-JUL-2001
DEFINITION  602760950F1 NIH_MGC_19 Homo sapiens cDNA clone IMAGE:4896324 5',
            mRNA sequence.
ACCESSION   BI199916
VERSION     BI199916.1  GI:14654937
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 699)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM1783  row: c  column: 13
            High quality sequence start: 5
            High quality sequence stop: 669.
FEATURES             Location/Qualifiers
     source          1. .699
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4896324"
                     /clone_lib="NIH_MGC_19"
                     /tissue_type="neuroblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: brain; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Library constructed  by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and  Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
BASE COUNT      120 a    216 c    222 g    141 t
ORIGIN


  Query Match           30.1%;  Score 30.4;  DB 10;  Length 699;
  Best Local Similarity   67.2%;  Pred. No. 78;
  Matches   43;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy      36 ggatgacgaactgcggaaacccggcgcccagatctacatgcgcttcatgcaggagcacac 95
           |||  ||||||| ||||  || ||  ||| || || |  | ||   ||||| | || |||
Db      78 GGACCACGAACTACGGACTCCAGGATCCCGGAACTCCGGGAGCGGGATGCACGCGCTCAC 19

Qy      96 ctgc 99
           || |
Db      18 CTCC 15


RESULT   7
BH140228/c
LOCUS       BH140228                 580 bp    DNA     linear   GSS 07-AUG-2001
DEFINITION  ZMMBBb0001L18f Maize B73 Zea mays genomic clone ZMMBBb0001L18f, DNA
            sequence.
ACCESSION   BH140228
VERSION     BH140228.1  GI:15099289
KEYWORDS    GSS.
SOURCE      Zea mays.
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 580)
  AUTHORS   Tomkins,J.P., Main,D., Goicoechea,J.L., Frisch,D.A. and Wing,R.A.
  TITLE     A Deep-Coverage BAC Library for Maize
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Wing RA
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 7288
            Fax: 864 656 4293
            Email: rwing@clemson.edu
            Seq primer: TAATACGACTCACTATAGGG
            Class: BAC ends
            High quality sequence start: 71
            High quality sequence stop: 510.
FEATURES             Location/Qualifiers
     source          1. .580
                     /organism="Zea mays"
                     /strain="B73"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBBb0001L18f"
                     /clone_lib="Maize B73"
                     /tissue_type="Young leaves"
                     /lab_host="E. coli"
                     /note="Vector: pCUGIBAC-1; Site_1: HindIII; Site_2: NotI;
                     For more details on library preparation, ordering clones
                     and sequence analysis see
                     http://www.genome.clemson.edu/projects/stc/maize/ZMMBBb "
BASE COUNT       74 a    221 c    138 g    146 t      1 others
ORIGIN


  Query Match           29.9%;  Score 30.2;  DB 12;  Length 580;
  Best Local Similarity   62.7%;  Pred. No. 85;
  Matches   47;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      27 gctgggctgggatgacgaactgcggaaacccggcgcccagatctacatgcgcttcatgca 86
           || | ||||||  ||| ||||| ||| || |||| |   |  || || ||||   | | |
Db     147 GCCGAGCTGGGCAGACAAACTGAGGAGACGCGGCTCGAGGGGCTCCACGCGCAGTAGGAA 88

Qy      87 ggagcacacctgcta 101
           |   |  ||| || |
Db      87 GTTTCCGACCGGCGA 73


RESULT   8
CNS01MA1/c
LOCUS       CNS01MA1                 894 bp    DNA     linear   GSS 14-JUN-2001
DEFINITION  Anopheles gambiae GSS SP6 end of clone 21E19 of NotreDame1 library
            from strain PEST of Anopheles gambiae (African malaria mosquito),
            genomic survey sequence.
ACCESSION   AL150666
VERSION     AL150666.1  GI:7011145
KEYWORDS    GSS.
SOURCE      African malaria mosquito.
  ORGANISM  Anopheles gambiae
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Nematocera;
            Culicoidea; Anopheles.
REFERENCE   1  (bases 1 to 894)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (16-FEB-2000) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
REFERENCE   2  (bases 1 to 894)
  AUTHORS   Roth,C.W., Brey,P.T., Ke,Z., Collins,F.H. and Weissenbach,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-FEB-2000) BBMI, Institut Pasteur, 25, rue du Dr.
            Roux, Paris 75015, France
COMMENT     This clone is from an A. gambiae BAC library provided by F.H.
            Collins and sequenced by Genoscope in collaboration with the
            Laboratory of Biochem. and Biol. Molec. of Insects, Institut
            Pasteur.
FEATURES             Location/Qualifiers
     source          1. .894
                     /organism="Anopheles gambiae"
                     /strain="PEST"
                     /db_xref="taxon:7165"
                     /clone="21E19"
                     /clone_lib="NotreDame1"
                     /note="end : SP6"
BASE COUNT      177 a    256 c    234 g    225 t      2 others
ORIGIN


  Query Match             29.9%;  Score 30.2;  DB 12;  Length 894;
  Best Local Similarity   62.7%;  Pred. No. 94;
  Matches   47;  Conservative    0;  Mismatches   28;  Indels    0;  Gaps    0;

Qy      27 gctgggctgggatgacgaactgcggaaacccggcgcccagatctacatgcgcttcatgca 86
           |||| |   |||| ||| ||||     || | || |||| || |||   | ||  ||||
Db     181 GCTGTGGGTGGATTACGTTCTGCTCGTGCCGGCCGACCAGTTCCACAACGGTTTGCTGCA 122

Qy      87 ggagcacacctgcta 101
           |||| | || | | |
Db     121 GGAGGAAACGTTCGA 107


RESULT   9
CNS02SYT
LOCUS       CNS02SYT                  910 bp    DNA     linear   GSS 15-MAY-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence T7 end of clone
            163G17 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL212510
VERSION     AL212510.1  GI:7871329
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1  (bases 1 to 910)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
  TITLE     Charaterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 910)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 910)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000) to the EMBL/GenBank/DDBJ databases
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .910
                     /organism="Tetraodon nigroviridis"
                     /db_xref="taxon:99883"
                     /clone="163G17"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0AG163AD09LP1-end : T7"
BASE COUNT      179 a    263 c    305 g    160 t      3 others
ORIGIN


  Query Match             29.5%;  Score 29.8;  DB 12;  Length 910;
  Best Local Similarity   66.2%;  Pred. No. 1.2e+02;
  Matches   43;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy      30 gggctgggatgacgaactgcggaaacccggcgcccagatctacatgcgcttcatgcagga 89
           ||||  | | | | | | || ||  ||| ||| || |  || ||| ||   |||||||||
Db     273 GGGCATGAAGGGCTACCAGCTGACCCCCAGCGACCTGGACTTCATCCGAGCCATGCAGGA 332

Qy      90 gcaca 94
           | |||
Db     333 GGACA 337


RESULT  10
CNS038S6
LOCUS       CNS038S6                  1010 bp    DNA     linear   GSS 15-MAY-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence T7 end of clone
            005N22 of library G from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL233007
VERSION     AL233007.1  GI:7892142
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1  (bases 1 to 1010)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
            Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
            Weissenbach,J.
  TITLE     Charaterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 1010)
  AUTHORS   Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Human gene number estimate provided by genome wide analysis using
            Tetraodon nigroviridis DNA sequence
  JOURNAL   Unpublished
REFERENCE   3  (bases 1 to 1010)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000) to the EMBL/GenBank/DDBJ databases
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1. .1010
                     /organism="Tetraodon nigroviridis"
                     /db_xref="taxon:99883"
                     /clone="005N22"
                     /clone_lib="G"
                     /note="Genoscope sequence ID : C0BG005DG11XD1-end : T7"
BASE COUNT      174 a    313 c    312 g    210 t      1 others
ORIGIN
```

```
  Query Match            29.5%;  Score 29.8;  DB 12;  Length 1010;
  Best Local Similarity  66.2%;  Pred. No. 1.3e+02;
  Matches   43;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy      30 gggctgggatgacgaactgcggaaacccggcgcccagatctacatgcgcttcatgcagga 89
           ||||  | | | | | | || ||  ||| ||| || |  || ||| ||   |||||||||
Db     620 GGGCATGAAGGGCTACCAGCTGACCCCCAGCGACCTGGACTTCATCCGAGCCATGCAGGA 679

Qy      90 gcaca 94
           | |||
Db     680 GGACA 684


RESULT  11
BM133818
LOCUS       BM133818                 655 bp    mRNA    linear   EST 07-JAN-2002
DEFINITION  NXLV_012_F09_F NXLV (Nsf Xylem Late wood Vertical) Pinus taeda cDNA
            clone NXLV_012_F09 5', mRNA sequence.
ACCESSION   BM133818
VERSION     BM133818.1  GI:17141057
KEYWORDS    EST.
SOURCE      loblolly pine.
  ORGANISM  Pinus taeda
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE   1  (bases 1 to 655)
  AUTHORS   Sederoff,R.
  TITLE     Molecular Basis of Wood Formation in the Pine Megagenome
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Johnson, Arthur
            North Carolina State University
            Tel: 919 515 7800
            Fax: 919 515 7801
            Email: ajohnson@unity.ncsu.edu
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .655
                     /organism="Pinus taeda"
                     /strain="Coastal plain loblolly pine from North Carolina"
                     /db_xref="taxon:3352"
                     /clone="NXLV_012_F09"
                     /clone_lib="NXLV (Nsf Xylem Late wood Vertical)"
                     /tissue_type="primary xylem"
                     /dev_stage="late wood"
                     /lab_host="XL1-Blue"
                     /note="Vector: pTriplEx; Site_1: EcoRI; The library is
                     from late (summer-August) wood, taken from below the crown
                     of a 20 year old tree. The harvested xylem tissue was on
                     the the cusp between transitional and mature wood. NOTE:
                     The sequences contain a 'cDNA adapter' between the EcoRI
                     site and the start of the EST. The adapter sequence is
                     'AATTCGGCCATTATGGCC'."
BASE COUNT      150 a    189 c    141 g    150 t     25 others
ORIGIN


  Query Match            29.3%;  Score 29.6;  DB 10;  Length 655;
  Best Local Similarity  57.6%;  Pred. No. 1.3e+02;
  Matches   53;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy       4 tccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcc 63
           | ||||    ||| ||||||  | |  || | | ||  | ||| || | | |||    |
Db     235 TTCCCGGGAACCCAATTTCCAGAACCAGGTTAGTATCCCCAACCGCCGCAGCCCTAATCA 294

Qy      64 cagatctacatgcgcttcatgcaggagcacac 95
           ||| ||  ||||| | ||    |    |||||
Db     295 CAGTTCATCATGCCCATCTACAATTTTCACAC 326


RESULT  12
BE269692
LOCUS       BE269692                 930 bp    mRNA    linear   EST 13-JUL-2000
DEFINITION  601185836F1 NIH_MGC_8 Homo sapiens cDNA clone IMAGE:3543584 5',
            mRNA sequence.
ACCESSION   BE269692
VERSION     BE269692.1  GI:9143317
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 930)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM237  row: g  column: 09
            High quality sequence stop: 238.
FEATURES             Location/Qualifiers
     source          1. .930
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3543584"
                     /clone_lib="NIH_MGC_8"
                     /tissue_type="Burkitt lymphoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lymph; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      226 a    228 c    267 g    209 t
ORIGIN


  Query Match            29.3%;  Score 29.6;  DB 10;  Length 930;
  Best Local Similarity  57.6%;  Pred. No. 1.4e+02;
  Matches   53;  Conservative    0;  Mismatches   39;  Indels    0;  Gaps    0;

Qy       7 ccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcccag 66
           || |   |||||||   | ||  ||||   | |||   |   || |||    |||| |||
Db     190 CCCCCCCCCCCATTGTACCAGGGGGGCGCAGGTGAGACAAAACGCAAAGAACGCGCACAG 249

Qy      67 atctacatgcgcttcatgcaggagcacacctg 98
           |    | | |||||  || || | |||  |||
Db     250 AAGCGCCTCCGCTTGGTGAAGCACCACCACTG 281


RESULT  13
BG394974
LOCUS       BG394974                 955 bp    mRNA    linear   EST 12-MAR-2001
DEFINITION  602457369F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:4579696 5',
            mRNA sequence.
ACCESSION   BG394974
VERSION     BG394974.1  GI:13288422
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 955)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
```

```
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM1297  row: j  column: 17
            High quality sequence stop: 717.
FEATURES             Location/Qualifiers
     source          1. .955
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4579696"
                     /clone_lib="NIH_MGC_16"
                     /tissue_type="retinoblastoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Library constructed  by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and  Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library."
BASE COUNT      268 a    208 c    225 g    254 t
ORIGIN


  Query Match             29.1%;  Score 29.4;  DB 10;  Length 955;
  Best Local Similarity   66.7%;  Pred. No. 1.6e+02;
  Matches   42;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy      33 ctgggatgacgaactgcggaaacccggcgcccagatctacatgcgcttcatgcaggagca 92
           ||  ||||||| ||| | ||||  | |||||  ||   | |  || |||||| |||   |
Db     836 CTTAGATGACGTACTCCCGAAAATCTGCGCCGCGAAAAAAACCCGTTTCATGAAGGTTGA 895

Qy      93 cac 95
           |||
Db     896 CAC 898


RESULT  14
BM468546
LOCUS       BM468546               1000 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6475521 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5578160
            5', mRNA sequence.
ACCESSION   BM468546
VERSION     BM468546.1  GI:18517588
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1000)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12332  row: m  column: 09
            High quality sequence stop: 725.
FEATURES             Location/Qualifiers
     source          1. .1000
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5578160"
                     /clone_lib="NIH_MGC_72"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      276 a    244 c    303 g    174 t      3 others
ORIGIN


  Query Match             28.9%;  Score 29.2;  DB 10;  Length 1000;
  Best Local Similarity   64.2%;  Pred. No. 1.9e+02;
  Matches   43;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy      22 cccaagctgggctgggatgacgaactgcggaaacccggcgcccagatctacatgcgcttc 81
           | |||| ||   |  || || || |||| ||| | | | |  |||| | | ||| ||| |
Db     690 CACAAGATGCTGTATGAAGAGGACCTGCAGAACCTCTGTGANCAGACCCAGATGAGCTCC 749

Qy      82 atgcagg 88
           |||||
Db     750 CAGCAGG 756


RESULT  15
AA264931
LOCUS       AA264931               313 bp    mRNA    linear   EST 19-APR-2001
DEFINITION  LD08544.5prime LD Drosophila melanogaster embryo BlueScript
            Drosophila melanogaster cDNA clone LD08544 5prime, mRNA sequence.
ACCESSION   AA264931
VERSION     AA264931.1  GI:1900974
KEYWORDS    EST.
SOURCE      fruit fly.
  ORGANISM  Drosophila melanogaster
            Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
            Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
            Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1  (bases 1 to 313)
  AUTHORS   Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
            Lewis,S. and Rubin,G.M.
  TITLE     BDGP/HHMI Drosophila EST Project
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Stapleton, M.
            BDGP
            Lawrence Berkeley National Lab
            One Cyclotron Rd, Berkeley, CA 94720, USA
            Fax: 510 486 6798
            Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
            Blast expect value = 1.1E-117 on X03121: Drosophila melanogaster
            serendipity (sry) locus DNA sequence
            Plate: 85  row: D  column: 8
            High quality sequence stop: 240.
FEATURES             Location/Qualifiers
     source          1. .313
                     /organism="Drosophila melanogaster"
                     /db_xref="BDGP_EST:BDcln007879"
                     /db_xref="taxon:7227"
                     /clone="LD08544"
                     /clone_lib="LD Drosophila melanogaster embryo BlueScript"
                     /sex="male and female"
                     /dev_stage="0 to 24 hours mixed stage embryonic"
                     /lab_host="SOLR"
                     /note="Organ: embryo; Vector: BlueScript SK; Site_1: EcoRI
                     ; Site_2: XhoI; Constructed using Stratagene ZAP-cDNA
                     Synthesis kit. Oligo dT-primed and directionally cloned at
                     EcoRI and XhoI in BlueScript SK(+/-)"
BASE COUNT       72 a     99 c     82 g     60 t
ORIGIN


  Query Match             28.7%;  Score 29;  DB 9;  Length 313;
  Best Local Similarity   71.7%;  Pred. No. 1.6e+02;
  Matches   38;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;
```

```
Qy      49 cggaaacccggcgcccagatctacatgcgcttcatgcaggagcacacctgcta 101
           | ||| |||||||||| | | | |  ||||| | | ||||||| | ||  ||| 
Db     235 CTGAAGCCCGGCGCCCGGCTGTGCCCGCGCTGCTTTCAGGAGCTCTCCGACTA 287
```

Search completed: October 3, 2002, 16:15:08
Job time: 16963 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 16:22:14 ; Search time 180.77 Seconds
(without alignments)
137.241 Million cell updates/sec

Title: US-09-826-581-5_COPY_500_600
Perfect score: 101
Sequence: 1 ctgtccccgcaggccccatt.....atgcaggagcacacctgcta 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	26.4	26.1	44377	2	US-08-804-227C-7	Sequence 7, Appli
	2	26.4	26.1	44377	2	US-08-804-198-1	Sequence 1, Appli
	3	26	25.7	4403765	4	US-09-103-840A-2	Sequence 2, Appli
	4	25.8	25.5	594	4	US-08-441-507-17	Sequence 17, Appl
	5	25.8	25.5	6085	4	US-09-029-603-4	Sequence 4, Appli
c	6	25.2	25.0	2980	2	US-08-463-081B-13	Sequence 13, Appl
c	7	25.2	25.0	2980	2	US-08-461-379A-13	Sequence 13, Appl
c	8	25.2	25.0	2980	2	US-08-462-390B-13	Sequence 13, Appl
c	9	25.2	25.0	2980	3	US-08-463-074B-13	Sequence 13, Appl
c	10	25.2	25.0	2980	3	US-08-465-585C-13	Sequence 13, Appl
c	11	25.2	25.0	2980	3	US-08-652-446-13	Sequence 13, Appl
	12	24.8	24.6	255	4	US-08-990-823-31	Sequence 31, Appl
c	13	24.8	24.6	1364	1	US-08-306-691B-50	Sequence 50, Appl
c	14	24.8	24.6	1364	5	PCT-US93-06251-65	Sequence 65, Appl
c	15	24.8	24.6	1900	4	US-09-608-285A-47	Sequence 47, Appl
c	16	24.8	24.6	2294	4	US-09-608-285A-49	Sequence 49, Appl
c	17	24.8	24.6	2371	4	US-09-608-285A-46	Sequence 46, Appl
c	18	24.8	24.6	2497	4	US-09-608-285A-51	Sequence 51, Appl
c	19	24.8	24.6	2693	4	US-09-608-285A-48	Sequence 48, Appl
c	20	24.8	24.6	2721	6	5215881-2	Patent No. 5215881
c	21	24.8	24.6	2762	4	US-09-608-285A-26	Sequence 26, Appl
c	22	24.8	24.6	2762	4	US-09-608-285A-52	Sequence 52, Appl
c	23	24.8	24.6	2762	4	US-09-240-639-1	Sequence 1, Appli
	24	24.8	24.6	3123	1	US-08-094-889-2	Sequence 2, Appli
c	25	24.8	24.6	8438	1	US-07-945-283-1	Sequence 1, Appli
	26	24.6	24.4	30001	1	US-08-125-468-1	Sequence 1, Appli
	27	24.6	24.4	30001	2	US-08-474-933-1	Sequence 1, Appli
	28	24.4	24.2	2327	4	US-08-868-435-1	Sequence 1, Appli
	29	24.4	24.2	2327	4	US-08-744-231-1	Sequence 1, Appli
	30	24.4	24.2	31571	1	US-08-323-443B-1	Sequence 1, Appli
	31	24.4	24.2	68750	3	US-09-335-409-1	Sequence 1, Appli
	32	24.4	24.2	68750	4	US-09-568-102-1	Sequence 1, Appli
	33	24.4	24.2	68750	4	US-09-567-969-1	Sequence 1, Appli
	34	24.4	24.2	68750	4	US-09-568-480-1	Sequence 1, Appli
	35	24.4	24.2	68750	4	US-09-568-486-1	Sequence 1, Appli
	36	24.4	24.2	68750	4	US-09-568-472-1	Sequence 1, Appli
	37	24.4	24.2	71989	4	US-09-443-501A-2	Sequence 2, Appli
	38	24.2	24.0	4695	2	US-08-231-193A-57	Sequence 57, Appl
	39	24.2	24.0	4695	2	US-08-486-273A-57	Sequence 57, Appl
	40	24.2	24.0	4695	3	US-08-940-086A-57	Sequence 57, Appl
	41	24.2	24.0	4695	4	US-08-940-035A-57	Sequence 57, Appl
	42	24	23.8	662	4	US-08-441-507-1	Sequence 1, Appli
	43	23.8	23.6	1406	4	US-09-287-097-1	Sequence 1, Appli
	44	23.8	23.6	3048	5	PCT-US95-14418-1	Sequence 1, Appli
	45	23.8	23.6	3048	5	PCT-US95-15327-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT   1
US-08-804-227C-7
; Sequence 7, Application US/08804227C
; Patent No. 5876991
;  GENERAL INFORMATION:
;    APPLICANT:  DeHoff, Bradley S.
;    APPLICANT:  Kuhstoss, Stuart A.
;    APPLICANT:  Rosteck, Paul R., Jr.
;    APPLICANT:  Sutton, Kimberly L.
;    TITLE OF INVENTION:  POLYKETIDE SYNTHASE GENES
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  THOMAS G. PLANT  1501
;      STREET:  LILLY CORPORATE CENTER
;      CITY:  INDIANAPOLIS
;      STATE:  IN
;      COUNTRY:  USA
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  ASCI(DOS) Text only
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/804,227C
;      FILING DATE:  February 21, 1997
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Plant, Thomas, G.
;      REGISTRATION NUMBER:  35,784
;      REFERENCE/DOCKET NUMBER:  X-8231
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-2459
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44377 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  350..14002
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  14046..20036
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  20110..31284
;    FEATURE:
```

```
;       NAME/KEY:  CDS
;       LOCATION:  31329..36071
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  36155..41830
US-08-804-227C-7


  Query Match             26.1%;  Score 26.4;  DB 2;  Length 44377;
  Best Local Similarity   69.2%;  Pred. No. 17;
  Matches   36;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      22 cccaagctgggctgggatgacgaactgcggaaacccggcgcccagatctaca 73
           ||   ||||||| | |  | |||||||| | |  |||| | || ||| ||||
Db   27484 CCGCTGCTGGGCAGCGGCGTCGAACTGCCGGAGTCCGGTGACCGGATGTACA 27535


RESULT   2
US-08-804-198-1
; Sequence 1, Application US/08804198
; Patent No. 5945320
;  GENERAL INFORMATION:
;    APPLICANT:  Burgett, Stanley G.
;    APPLICANT:  Kuhstoss, Stuart A.
;    APPLICANT:  Rao, Nagaraja R.
;    APPLICANT:  Richardson, Mark A.
;    APPLICANT:  Rosteck, Paul R., Jr.
;    TITLE OF INVENTION:  PLATENOLIDE SYNTHASE GENE
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PAUL R. CANTRELL  1138
;      STREET:  LILLY CORPORATE CENTER
;      CITY:  INDIANAPOLIS
;      STATE:  IN
;      COUNTRY:  USA
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  Macintosh
;      OPERATING SYSTEM:  Macintosh 7.0
;      SOFTWARE:  Microsoft Word 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/804,198
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  CANTRELL, PAUL R.
;      REGISTRATION NUMBER:  36,470
;      REFERENCE/DOCKET NUMBER:  P9113
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3885
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  44377 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  350..14002
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  14046..20036
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  20110..31284
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  31329..36071
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  36155..41830
US-08-804-198-1


  Query Match             26.1%;  Score 26.4;  DB 2;  Length 44377;
  Best Local Similarity   69.2%;  Pred. No. 17;
  Matches   36;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy      22 cccaagctgggctgggatgacgaactgcggaaacccggcgcccagatctaca 73
           ||   ||||||| | |  | |||||||| | |  |||| | || ||| ||||
Db   27484 CCGCTGCTGGGCAGCGGCGTCGAACTGCCGGAGTCCGGTGACCGGATGTACA 27535


RESULT   3
US-09-103-840A-2
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
;  APPLICANT: FLEISCHMAN, Robert D.
;  APPLICANT:  WHITE, Owen R.
;  APPLICANT:  FRASER, Claire M.
;  APPLICANT:  VENTER, John C.
;  TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
;  TITLE OF INVENTION:  TUBERCULOSIS
;  FILE REFERENCE: 24366-20007.00
;  CURRENT APPLICATION NUMBER: US/09/103,840A
;  CURRENT FILING DATE:  1998-06-24
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 4403765
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   FEATURE:
;   OTHER INFORMATION: CDC 1551
;   OTHER INFORMATION: "n" bases at various positions throughout the sequence
;   OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2


  Query Match             25.7%;  Score 26;  DB 4;  Length 4403765;
  Best Local Similarity   58.8%;  Pred. No. 40;
  Matches   47;  Conservative    0;  Mismatches   33;  Indels    0;  Gaps    0;

Qy       6 cccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgccca 65
           |||||| ||     | | ||| | ||    || ||||||  |  || | |||||| | ||
Db 2113570 cccgcacgcgtggatccgcaacccggttacggctgacgagatctggcagcccggcccaca 2113629

Qy      66 gatctacatgcgcttcatgc 85
           || |   |||  |  |  ||
Db 2113630 gaaccggatggtcagctggc 2113649


RESULT   4
US-08-441-507-17
; Sequence 17, Application US/08441507
; Patent No. 6214358
;  GENERAL INFORMATION:
;    APPLICANT:  Singh, Mohan Bir;
;    APPLICANT:  Smith, Penelope; and
;    APPLICANT:  Knox, Robert Bruce
;    TITLE OF INVENTION:  Protein Allergens of the Species Cynodon
;    TITLE OF INVENTION:  Dactylon
;    NUMBER OF SEQUENCES:  52
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  LAHIVE & COCKFIELD LLP
;      STREET:  28 State Street
;      CITY:  Boston
;      STATE:  Massachusetts
;      COUNTRY:  USA
;      ZIP:  02109
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  ASCII text
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/441,507
;       FILING DATE:  15-May-1995
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/969,875
;       FILING DATE:  30-October-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Mandragouras, Amy E.
;       REGISTRATION NUMBER:  36,207
;       REFERENCE/DOCKET NUMBER:  IMI-049DV
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 227-7400
;       TELEFAX:  (617) 742-4214
;   INFORMATION FOR SEQ ID NO:  17:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  594 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
US-08-441-507-17


  Query Match             25.5%;  Score 25.8;  DB 4;  Length 594;
  Best Local Similarity   63.9%;  Pred. No. 9.7;
  Matches   39;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       8 cgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcccaga 67
           ||   | | |  | |||||  ||||  ||| | ||||| ||  ||||||  | ||
Db      17 CGTTCGGCGCCATGGCCAAGAAGGGCGAGGAGGACAAGCTGCGCAAGGCCGGCGAGCTGA 76

Qy      68 t 68
           |
Db      77 T 77


RESULT   5
US-09-029-603-4
; Sequence 4, Application US/09029603
; Patent No. 6210935
; GENERAL INFORMATION:
;  APPLICANT: Schupp, Thomas
;  APPLICANT: Engel, Natalie
;  APPLICANT: Bietenhader, Jurg
;  APPLICANT: Toupet, Christine
;  APPLICANT: Pospiech, Andreas
;  TITLE OF INVENTION: Staurosporin Biosynthesis Gene Clusters
;  FILE REFERENCE: 4-20555/A/PCT
;  CURRENT APPLICATION NUMBER: US/09/029,603
;  CURRENT FILING DATE: 1998-03-20
;  EARLIER APPLICATION NUMBER: PCT/EP96/03643
;  EARLIER FILING DATE: 1996-08-19
;  NUMBER OF SEQ ID NOS: 11
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
;   LENGTH: 6085
;   TYPE: DNA
;   ORGANISM: Streptomyces longisporoflavus
;   FEATURE:
;   NAME/KEY: misc_RNA
;   LOCATION: (378)..(1665)
;   OTHER INFORMATION: ORF
;   FEATURE:
;   NAME/KEY: misc_RNA
;   LOCATION: (1747)..(2553)
;   OTHER INFORMATION: ORF
;   FEATURE:
;   NAME/KEY: misc_RNA
;   LOCATION: (2593)..(4011)
;   OTHER INFORMATION: ORF
;   FEATURE:
;   NAME/KEY: misc_RNA
;   LOCATION: (4013)..(4999)
;   OTHER INFORMATION: ORF
;   FEATURE:
;   NAME/KEY: misc_RNA
;   LOCATION: (5071)..(6085)
;   OTHER INFORMATION: ORF
US-09-029-603-4


  Query Match             25.5%;  Score 25.8;  DB 4;  Length 6085;
  Best Local Similarity   56.5%;  Pred. No. 17;
  Matches   48;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy      15 cccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcccagatctacat 74
           ||  |  | ||  || |||  |   | ||||||| | |    | || || | || ||||
Db    3243 ccgctggctcaccctcggccagatcgccgaactgctgcacgaggacgacctggtcaacat 3302

Qy      75 gcgcttcatgcaggagcacacctgc 99
             |  || |  || ||   | |||
Db    3303 gaacgccaggacggtgctgtcgtgc 3327


RESULT   6
US-08-463-081B-13/c
; Sequence 13, Application US/08463081B
; Patent No. 5871960
; Patent No. 5871960 5837487
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, Kendall A. & Beadling, Carol
;    TITLE OF INVENTION:  Nucleic Acids Encoding CR5 Polypeptide,
;    TITLE OF INVENTION:  Vector and Transformed Cell Thereof, and Expression Thereo
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PRETTY, SCHROEDER & POPLAWSKI
;      STREET:  444 South Flower St. - Suite 1900
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0,
;      SOFTWARE:  Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,081B
;      FILING DATE:  5-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/104,736
;      FILING DATE:  10-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/796,066
;      FILING DATE:  20-NOV-91
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Viviana Amzel, Ph. D.
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  P66 38150 (DART-060)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 622-7700
;      TELEFAX:  (213) 489-4210
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2980 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  240..1475
```

```
US-08-463-081B-13


  Query Match             25.0%;  Score 25.2;  DB 2;  Length 2980;
  Best Local Similarity   71.7%;  Pred. No. 22;
  Matches   33;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      20 ttcccaagctgggctgggatgacgaactgcggaaacccggcgcccca 65
           |||| | || ||| ||||  |  || ||||||||||||  | || |
Db    2119 TTCCGAGGCAGGGGTGGGGAGCAGACCTGCGGAAACCCTCCCCCAA 2074


RESULT   7
US-08-461-379A-13/c
; Sequence 13, Application US/08461379A
; Patent No. 5871961
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, Kendall A. & Beadling, Carol
;    TITLE OF INVENTION:  Nucleic Acids Encoding CR5 Polypeptide,
;    TITLE OF INVENTION:  Vector and Transformed Cell Thereof, and
;    TITLE OF INVENTION:  Expression Thereof
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ratner & Prestia
;      ADDRESSEE:  .          (B) STREET:One Westlakes-Berwyn
;      CITY:  Valley Forge
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19482
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0,
;      SOFTWARE:  Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/461,379A
;      FILING DATE:  5-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/330,108; 08/104,736
;      APPLICATION NUMBER:  & 07/796,066
;      FILING DATE:  27-OCT-1994; 10-AUG-1993 & 20-NOV-91
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Viviana Amzel, Ph. D.
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  DART-070
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610)470-0700
;      TELEFAX:  (610)470-0701
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2980 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  240..1475
US-08-461-379A-13


  Query Match             25.0%;  Score 25.2;  DB 2;  Length 2980;
  Best Local Similarity   71.7%;  Pred. No. 22;
  Matches   33;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      20 ttcccaagctgggctgggatgacgaactgcggaaacccggcgcccca 65
           |||| | || ||| ||||  |  || ||||||||||||  | || |
Db    2119 TTCCGAGGCAGGGGTGGGGAGCAGACCTGCGGAAACCCTCCCCCAA 2074


RESULT   8
US-08-462-390B-13/c
; Sequence 13, Application US/08462390B
; Patent No. 5882894
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, K. A., & Beadling, C.
;    TITLE OF INVENTION:  Nucleic Acids Encoding CR8 Polypeptide, Vector and
;    TITLE OF INVENTION:  Transformed Cell Thereof, and Expression Thereof
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Ratner & Prestia
;      ADDRESSEE:  .          (B) STREET:One Westlakes-Berwyn
;      CITY:  Valley Forge
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19482
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/462,390B
;      FILING DATE:  5-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/330,108
;      FILING DATE:  27-OCT-1994
;      APPLICATION NUMBER:  USSN 08/104,736
;      FILING DATE:  10-AUG-1993
;      APPLICATION NUMBER:  USSN 07/796,066
;      FILING DATE:  20-NOV-91
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Viviana Amzel, Ph. D.
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  DART-040
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (610)407-0700
;      TELEFAX:  (610)407-0701
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2980 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  240..1475
US-08-462-390B-13


  Query Match             25.0%;  Score 25.2;  DB 2;  Length 2980;
  Best Local Similarity   71.7%;  Pred. No. 22;
  Matches   33;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      20 ttcccaagctgggctgggatgacgaactgcggaaacccggcgcccca 65
           |||| | || ||| ||||  |  || ||||||||||||  | || |
Db    2119 TTCCGAGGCAGGGGTGGGGAGCAGACCTGCGGAAACCCTCCCCCAA 2074


RESULT   9
US-08-463-074B-13/c
; Sequence 13, Application US/08463074B
; Patent No. 6020155
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, Kendall A. & Beadling, Carol
;    TITLE OF INVENTION:  Nucleic Acids Encoding CR1 Fusion Protein, Vector an
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PRETTY, SCHROEDER & POPLAWSKI
;      ADDRESSEE:  .          (B) STREET:                444 South Flower St. - Suite 1
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0,
;      SOFTWARE:  Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/463,074B
;      FILING DATE:  5-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/104,736
;      FILING DATE:  10-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/796,066
;      FILING DATE:  20-NOV-91
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Viviana Amzel, Ph. D.
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  P66 38143 (DART-020)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 622-7700
;      TELEFAX:  (213) 489-4210
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2980 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  240..1475
US-08-463-074B-13


  Query Match             25.0%;  Score 25.2;  DB 3;  Length 2980;
  Best Local Similarity   71.7%;  Pred. No. 22;
  Matches   33;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      20 ttcccaagctgggctgggatgacgaactgcggaaacccggcgcccca 65
           |||| | || ||| ||||  |  || ||||||||||||  | || |
Db    2119 TTCCGAGGCAGGGGTGGGGAGCAGACCTGCGGAAACCCTCCCCCAA 2074


RESULT  10
US-08-465-585C-13/c
; Sequence 13, Application US/08465585C
; Patent No. 6027914
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, K. A., & Beadling, C.
;    TITLE OF INVENTION:  Nucleic Acids Encoding CR6 Polypeptide, Vecto
;    TITLE OF INVENTION:  Transformed Cell Thereof, and Expression Thereof
;    NUMBER OF SEQUENCES:  35
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PRETTY,SCHROEDER & POPLAWSKI
;      ADDRESSEE:  .          (B) STREET:                  444South Flower St. - Suite 190
;      CITY:  Los Angeles
;      STATE:  Californiaa
;      COUNTRY:  USA
;      ZIP:  900071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/465,585C
;      FILING DATE:  5-JUNE-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  USSN 08/330,108
;      FILING DATE:  27-OCT-1994
;      APPLICATION NUMBER:  USSN 08/104,736
;      FILING DATE:  10-AUG-1993
;      APPLICATION NUMBER:  USSN 07/796,066
;      FILING DATE:  20-NOV-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Viviana Amzel, Ph. D.
;      REGISTRATION NUMBER:  30,930
;      REFERENCE/DOCKET NUMBER:  P66 38149 (DART-050)
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (213) 622-7700
;      TELEFAX:  (213) 4894210
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2980 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  240..1475
US-08-465-585C-13


  Query Match             25.0%;  Score 25.2;  DB 3;  Length 2980;
  Best Local Similarity   71.7%;  Pred. No. 22;
  Matches   33;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      20 ttcccaagctgggctgggatgacgaactgcggaaacccggcgcccca 65
           |||| | || ||| ||||  |  || ||||||||||||  | || |
Db    2119 TTCCGAGGCAGGGGTGGGGAGCAGACCTGCGGAAACCCTCCCCCAA 2074


RESULT  11
US-08-652-446-13/c
; Sequence 13, Application US/08652446
; Patent No. 6057427
;  GENERAL INFORMATION:
;    APPLICANT:  Smith, Kendall A. & Beadling, Carol
;    TITLE OF INVENTION:  Nucleic Acids Encoding CR5
;    TITLE OF INVENTION:  Polypeptide, Vector and Transformed Cell Thereof, and
;    TITLE OF INVENTION:  Expression Thereof
;    NUMBER OF SEQUENCES:  38
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  PRETTY, SCHROEDER & POPLAWSKI
;      ADDRESSEE:  .          (B) STREET:                  444 South Flower St. - Suite
;      CITY:  Los Angeles
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  90071
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0,
;      SOFTWARE:  Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/652,446
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  EP App. # 96921319.8
;      FILING DATE:  5-JAN-1998
;      APPLICATION NUMBER:  PCT/US/96/09194
;      FILING DATE:  5-JUN-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/330,108
;      FILING DATE:  27-OCT-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/463,074
;      FILING DATE:  5-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/462,337
;      FILING DATE:  5-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/462,390
```

```
;       FILING DATE:  5-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/465,585
;       FILING DATE:  5-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/463,081
;       FILING DATE:  5-JUN-1995
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/461,379
;       FILING DATE:  5-JUN-1991
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/739,523
;       FILING DATE:  29-OCT-1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Viviana Amzel, Ph. D.
;       REGISTRATION NUMBER:  30,930
;       REFERENCE/DOCKET NUMBER:  FP66 40035
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (213) 622-7700
;       TELEFAX:  (213) 489-4210
;   INFORMATION FOR SEQ ID NO:  13:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2980 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  240..1475
US-08-652-446-13


  Query Match             25.0%;  Score 25.2;  DB 3;  Length 2980;
  Best Local Similarity   71.7%;  Pred. No. 22;
  Matches   33;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy      20 ttcccaagctgggctgggatgacgaactgcggaaacccggcgccca 65
           |||| | || ||| ||||  |  || ||||||||||||  | || |
Db    2119 TTCCGAGGCAGGGGTGGGGAGCAGACCTGCGGAAACCCTCCCCCAA 2074


RESULT  12
US-08-990-823-31
; Sequence 31, Application US/08990823D
; Patent No. 6228371
; GENERAL INFORMATION:
;  APPLICANT: Nano, Francis
;  TITLE OF INVENTION: Mycobacterium Tuberculosis DNA Sequences Encoding
;  TITLE OF INVENTION: immunostimulatory Peptides
;  FILE REFERENCE: 49086
;  CURRENT APPLICATION NUMBER: US/08/990,823D
;  CURRENT FILING DATE: 1997-12-15
;  EARLIER APPLICATION NUMBER: US 96/10375
;  EARLIER FILING DATE: 1996-06-14
;  EARLIER APPLICATION NUMBER: 60/000,254
;  EARLIER FILING DATE: 1995-06-15
;  NUMBER OF SEQ ID NOS: 113
;  SOFTWARE: PatentIN Ver. 2.0
; SEQ ID NO 31
;   LENGTH: 255
;   TYPE: DNA
;   ORGANISM: Mycobacterium tuberculosis
;   FEATURE:
;   NAME/KEY: Modified base
;   OTHER INFORMATION: n represents a or g or c or t/u
US-08-990-823-31


  Query Match             24.6%;  Score 24.8;  DB 4;  Length 255;
  Best Local Similarity   53.0%;  Pred. No. 17;
  Matches   53;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy       2 tgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggcg 61
           |||||||||| | |  |  |  ||||||  | |   ||   |   |  | | |    ||||
Db     143 tgtccccgctgccttcgctggccaagcgcgccgccgacatcctgcgccaggacgtgggcg 202

Qy      62 cccagatctacatgcgcttcatgcaggagcacacctgcta 101
             ||| || ||||   |    |  | | || ||||  | |
Db     203 ggcaggtcgacattttcgatgtcaatgtgcccaccatcca 242


RESULT  13
US-08-306-691B-50/c
; Sequence 50, Application US/08306691B
; Patent No. 5734039
;  GENERAL INFORMATION:
;    APPLICANT:  Calabretta, Bruno
;    APPLICANT:  Skorski, Tomasz
;    TITLE OF INVENTION:  ANTISENSE
;    TITLE OF INVENTION:  OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
;    NUMBER OF SEQUENCES:  55
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seidel, Gonda, Lavorgna & Monaco, P.C.
;      STREET:  Two Penn Center, Suite 1800
;      CITY:  Philadelphia
;      STATE:  Pennsylvania
;      COUNTRY:  U.S.A.
;      ZIP:  19102
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette, 3.50 inch, 720 Kb
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  MS-DOS
;      SOFTWARE:  WordPerfect 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/306,691B
;      FILING DATE:  September 15, 1994
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Monaco, Daniel A.
;      REGISTRATION NUMBER:  30,480
;      REFERENCE/DOCKET NUMBER:  8321-8
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (215) 568-8383
;      TELEFAX:  (215) 568-5549
;      TELEX:  No. 5734039e
;  INFORMATION FOR SEQ ID NO:  50:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1364 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-306-691B-50


  Query Match             24.6%;  Score 24.8;  DB 1;  Length 1364;
  Best Local Similarity   54.3%;  Pred. No. 25;
  Matches   50;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy       9 gcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcccagat 68
           | || ||   |  || |  ||||| |||| || |    || ||  | |  |  |    ||
Db     529 GGAGACCTGGTGGCCAAGACTGGGATGGGGTGGCACCATGGGGGTATCGAGGACGTGCAT 470

Qy      69 ctacatgcgcttcatgcaggagcacacctgct 100
           || |    ||| ||||  |  | | | ||||| 
Db     469 CTGCTCCAGCTCCATGTGGCGGTAGAGCTGCT 438


RESULT  14
PCT-US93-06251-65/c
; Sequence 65, Application PC/TUS9306251
;  GENERAL INFORMATION:
;    APPLICANT:  Wickstrom, Eric and Rife, Jason P.
```

```
;    TITLE OF INVENTION:  Trivalent Synthesis of Oligonucleotides Containing
;    TITLE OF INVENTION:  Stereospecific Alkylphosphonates and Arylphosphonates
;    NUMBER OF SEQUENCES:  93
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SCULLY, SCOTT, MURPHY & PRESSER
;      STREET:  400 Garden City Plaza
;      CITY:  Garden City
;      STATE:  NY
;      COUNTRY:  USA
;      ZIP:  11530
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/06251
;      FILING DATE:  19930630
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DiGiglio, Frank S.
;      REGISTRATION NUMBER:  31,346
;      REFERENCE/DOCKET NUMBER:  8586
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  516-742-4343
;      TELEFAX:  516-742-4366
;      TELEX:  230 901 SANS UR
;  INFORMATION FOR SEQ ID NO:  65:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1364 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
PCT-US93-06251-65


  Query Match             24.6%;  Score 24.8;  DB 5;  Length 1364;
  Best Local Similarity   54.3%;  Pred. No. 25;
  Matches   50;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy       9 gcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggcgcccagat 68
           | || ||   |  || |  ||||| |||| || |    || ||  | |  |  |    ||
Db     529 GGAGACCTGGTGGCCAAGACTGGGATGGGGTGGCACCATGGGGGTATCGAGGACGTGCAT 470

Qy      69 ctacatgcgcttcatgcaggagcacacctgct 100
             || |    ||| ||||  |  | | | |||||
Db     469 CTGCTCCAGCTCCATGTGGCGGTAGAGCTGCT 438


RESULT  15
US-09-608-285A-47/c
; Sequence 47, Application US/09608285A
; Patent No. 6335013
; GENERAL INFORMATION:
;  APPLICANT:  Ford, John
;  APPLICANT:  Mulero, Julio
;  APPLICANT:  Yeung, George
;  TITLE OF INVENTION:  METHODS AND MATERIALS RELATING TO CD39-LIKE
;  TITLE OF INVENTION:  POLYPEPTIDES
;  FILE REFERENCE:  28110/36570
;  CURRENT APPLICATION NUMBER: US/09/608,285A
;  CURRENT FILING DATE:  2000-06-30
;  PRIOR APPLICATION NUMBER:  09/583,231
;  PRIOR FILING DATE:  2000-05-26
;  PRIOR APPLICATION NUMBER:  09/557,800
;  PRIOR FILING DATE:  2000-04-25
;  PRIOR APPLICATION NUMBER:  09/481,238
;  PRIOR FILING DATE:  2000-01-11
;  PRIOR APPLICATION NUMBER:  09/370,265
;  PRIOR FILING DATE:  1999-08-09
;  PRIOR APPLICATION NUMBER:  PCT/US99/16180
;  PRIOR FILING DATE:  1999-07-16
;  PRIOR APPLICATION NUMBER:  09/350,836
;  PRIOR FILING DATE:  1999-07-09
;  PRIOR APPLICATION NUMBER:  09/273,447
;  PRIOR FILING DATE:  1999-03-19
;  PRIOR APPLICATION NUMBER:  09/244,444
;  PRIOR FILING DATE:  1999-02-04
;  PRIOR APPLICATION NUMBER:  09/122,449
;  PRIOR FILING DATE:  1998-07-24
;  PRIOR APPLICATION NUMBER:  09/118,205
;  PRIOR FILING DATE:  1998-07-16
;  NUMBER OF SEQ ID NOS:  60
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 47
;   LENGTH: 1900
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-608-285A-47


  Query Match             24.6%;  Score 24.8;  DB 4;  Length 1900;
  Best Local Similarity   53.0%;  Pred. No. 27;
  Matches   53;  Conservative    0;  Mismatches   47;  Indels    0;  Gaps    0;

Qy       1 ctgtccccgcaggccccatttcccaagctgggctgggatgacgaactgcggaaacccggc 60
           ||||  |  ||||  ||    | |   || ||||| |||  |||| | |   | | |
Db     551 CTGTGTCTCCAGGGTCCGACACACGTACTTGGCTGCGATCTCGAAGTCCCCCACCACCAG 492

Qy      61 gcccagatctacatgcgcttcatgcaggagcacacctgct 100
           ||      |   | ||| ||    |||  |||||| |||
Db     491 GCTGCCTCCCTTCTCCGCATCTATGAGGCCCACACCAGCT 452


Search completed: October  3, 2002, 16:24:51
Job time: 17131 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 14:49:54 ; Search time 7316.32 Seconds
(without alignments)
288.886 Million cell updates/sec

Title: US-09-826-581-5_COPY_1000_1100
Perfect score: 101
Sequence: 1 gctcaagttcctgcacatct.....cacattccgagacttggctg 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	101	100.0	1647	6	AX281582	AX281582 Sequence
2	101	100.0	2109	6	AX099776	AX099776 Sequence
3	101	100.0	2115	6	AX099802	AX099802 Sequence
4	101	100.0	2115	9	AF214519	AF214519 Homo sapi
5	101	100.0	2290	9	HSA249977	AJ249977 Homo sapi
6	93.6	92.7	1867	6	AX099774	AX099774 Sequence
7	93.6	92.7	1873	4	AF214520	AF214520 Sus scrof
8	93.6	92.7	1873	6	AX099800	AX099800 Sequence
9	93.6	92.7	2022	6	AX099804	AX099804 Sequence
10	79.2	78.4	1722	6	AX281580	AX281580 Sequence
c 11	79.2	78.4	152129	2	AC027416	AC027416 Homo sapi
c 12	79.2	78.4	196554	2	AC073128	AC073128 Homo sapi
c 13	79.2	78.4	206854	9	AC009974	AC009974 Homo sapi
14	76	75.2	5888	4	AF214521	AF214521 Sus scrof
c 15	71	70.3	227724	2	AF336381	AF336381 Mus muscu
c 16	32.8	32.5	159891	9	AC022469	AC022469 Homo sapi
17	31.6	31.3	221694	2	AC095460	AC095460 Rattus no
18	31.2	30.9	225392	2	AC095006	AC095006 Rattus no
19	30.8	30.5	1167	9	AB025580	AB025580 Homo sapi
20	30.8	30.5	1435	6	AR139104	AR139104 Sequence
21	30.8	30.5	2194	9	AF087875	AF087875 Homo sapi
22	30.8	30.5	2203	9	BC020540	BC020540 Homo sapi
23	30.8	30.5	2223	9	AK001887	AK001887 Homo sapi
24	30.8	30.5	3132	10	BC015283	BC015283 Mus muscu
25	30.8	30.5	65201	9	HS754E20A	AL096761 Human DNA
c 26	30.8	30.5	122732	2	HS436C18	AL031301 Homo sapi
27	30.4	30.1	166703	2	AC096427	AC096427 Rattus no
28	30.4	30.1	187787	2	AC094522	AC094522 Rattus no
29	30.2	29.9	71517	9	AC073862	AC073862 Homo sapi
30	30.2	29.9	313050	1	AJ414152	AJ414152 Yersinia
31	30	29.7	106	11	G67375	G67375 A898 Human
c 32	30	29.7	64773	2	AC091131	AC091131 Homo sapi
33	29.8	29.5	1681	4	AF334948	AF334948 Canis fam
34	29.8	29.5	34815	2	AC090221	AC090221 Homo sapi
c 35	29.8	29.5	75131	2	AC023022	AC023022 Homo sapi
36	29.8	29.5	81767	2	AC021929	AC021929 Homo sapi
37	29.8	29.5	110000	2	CEY105E8_2	Continuation (3 of
38	29.8	29.5	153170	9	HS1103G7	AL034548 Human DNA
39	29.8	29.5	171747	9	AP001554	AP001554 Homo sapi
c 40	29.8	29.5	173071	9	AC009477	AC009477 Homo sapi
41	29.8	29.5	184926	9	AP000757	AP000757 Homo sapi
42	29.8	29.5	206463	2	AP000834	AP000834 Homo sapi
43	29.8	29.5	277607	3	CEY105E8A	AL132876 Caenorhab
c 44	29.6	29.3	566	6	AX154810	AX154810 Sequence
45	29.6	29.3	1358	8	AY065231	AY065231 Arabidops

ALIGNMENTS

RESULT 1
AX281582
LOCUS AX281582 1647 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 5 from Patent WO0177305.
ACCESSION AX281582
VERSION AX281582.1 GI:16608833
KEYWORDS .
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (sites)
AUTHORS Andersson,L., Luthman,H. and Marklund,S.
TITLE Variants of the human amp-activated protein kinase gamma 3 subunit
JOURNAL Patent: WO 0177305-A 5 18-OCT-2001;
Arexis AB (SE)
FEATURES Location/Qualifiers
source 1. .1647
/organism="Homo sapiens"
/db_xref="taxon:9606"
CDS 20. .1489
/note="unnamed protein product"
/codon_start=1

```
                     /protein_id="CAD10589.1"
                     /db_xref="GI:16608834"
                     /translation="MEPGLEHALRRTPSWSSLGGSEHQEMSFLEQENSSSWPSPAVTS
                     SSERIRGKRRAKALRWTRQKSVEEGEPPGQGEGPRSRPAAESTGLEATFPKTTPLAQA
                     DPAGVGTPPTGWDCLPSDCTASAAGSSTDDVELATEFPATEAWECELEGLLEERPALC
                     LSPQAPFPKLGWDDELRKPGAQIYMRFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFA
                     LVANGVRAAPLWDSKKQSFVGMLTITDFILVLHRYYRSPLVQIYEIEQHKIETWREIY
                     LQGCFKPLVSISPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIF
                     GSLLPRPSFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSALPVVNECGQV
                     VGLYSRFDVIHLAAQQTYNHLDMSVGEALRQRTLCLEGVLSCQPHESLGEVIDRIARE
                     QVHRLVLVDETQHLLGVVSLSDILQALVLSPAGIDALGA"
BASE COUNT      346 a    502 c    462 g    337 t
ORIGIN


  Query Match             100.0%;  Score 101;  DB 6;  Length 1647;
  Best Local Similarity   100.0%;  Pred. No. 1e-20;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 gctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcac 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1000 GCTCAAGTTCCTGCACATCTTTGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCAC 1059

Qy     61 tatccaagatttgggcatcggcacattccgagacttggctg 101
          |||||||||||||||||||||||||||||||||||||||||
Db   1060 TATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCTG 1100


RESULT   2
AX099776
LOCUS       AX099776                2109 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 3 from Patent WO0120003.
ACCESSION   AX099776
VERSION     AX099776.1  GI:13538810
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2109)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 3 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .2109
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             472. .1389
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35799.1"
                     /db_xref="GI:13538811"
                     /translation="MRFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAP
                     LWDSKKQSFVGMLTITDFILVLHRYYRSPLVQIYEIEQHKIETWREIYLQGCFKPLVS
                     ISPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLPRPSFL
                     YRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSALPVVNECGQVVGLYSRFDVI
                     HLAAQQTYNHLDMSVGEALRQRTLCLEGVLSCQPHESLGEVIDRIAREQVHRLVLVDE
                     TQHLLGVVSLSDILQALVLSPAGIDALGA"
BASE COUNT      458 a    621 c    560 g    470 t
ORIGIN


  Query Match             100.0%;  Score 101;  DB 6;  Length 2109;
  Best Local Similarity   100.0%;  Pred. No. 1e-20;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 gctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcac 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    900 GCTCAAGTTCCTGCACATCTTTGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCAC 959

Qy     61 tatccaagatttgggcatcggcacattccgagacttggctg 101
          |||||||||||||||||||||||||||||||||||||||||
Db    960 TATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCTG 1000


RESULT   3
AX099802
LOCUS       AX099802                2115 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 29 from Patent WO0120003.
ACCESSION   AX099802
VERSION     AX099802.1  GI:13538836
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2115)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 29 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .2115
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             1. .1395
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35801.1"
                     /db_xref="GI:13538837"
                     /translation="MSFLEQENSSSWPSPAVTSSSERIRGKRRAKALRWTRQKSVEEG
                     EPPGQGEGPRSRPTAESTGLEATFPKTTPLAQADPAGVGTPPTGWDCLPSDCTASAAG
                     SSTDDVELATEFPATEAWECELEGLLEERPALCLSPQAPFPKLGWDDELRKPGAQIYM
                     RFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEQHKIETWREIYLQGCFKPLVSISPNDSLFEAVYTLI
                     KNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNECGQVVGLYSRFDVIHLAAQQTYNHLDMSV
                     GEALRQRTLCLEGVLSCQPHESLGEVIDRIAREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT      460 a    622 c    562 g    471 t
ORIGIN


  Query Match             100.0%;  Score 101;  DB 6;  Length 2115;
  Best Local Similarity   100.0%;  Pred. No. 1e-20;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 gctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcac 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    906 GCTCAAGTTCCTGCACATCTTTGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCAC 965

Qy     61 tatccaagatttgggcatcggcacattccgagacttggctg 101
          |||||||||||||||||||||||||||||||||||||||||
Db    966 TATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCTG 1006


RESULT   4
AF214519
LOCUS       AF214519                2115 bp    mRNA    linear   PRI 03-JUN-2000
DEFINITION  Homo sapiens AMP-activated protein kinase gamma subunit (PRKAG3)
            mRNA, complete cds.
ACCESSION   AF214519
VERSION     AF214519.1  GI:8215681
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2115)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A., Thelander,M.,
            Rogel-Gaillard,C., Paul,S., Iannuccelli,N., Rask,L., Ronne,H.,
            Lundstrom,K., Reinsch,N., Gellin,J., Kalm,E., Roy,P.L., Chardon,P.
            and Andersson,L.
  TITLE     A mutation in PRKAG3 associated with excess glycogen content in pig
            skeletal muscle
  JOURNAL   Science 288 (5469), 1248-1251 (2000)
  MEDLINE   20280150
   PUBMED   10818001
REFERENCE   2  (bases 1 to 2115)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A.,
            Rogel-Gaillard,C., Paul,S., Gellin,J., Lundstrom,K., Reinsch,N.,
            Kalm,E., Le Roy,P., Chardon,P. and Andersson,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-1999) Dept Animal Breeding and Genetics, Swedish
            University of Agricultural Sciences, BMC box 597, Uppsala 751 24,
            Sweden
FEATURES             Location/Qualifiers
     source          1. .2115
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2p"
                     /tissue_type="skeletal muscle"
     gene            1. .2115
                     /gene="PRKAG3"
     CDS             1. .1395
                     /gene="PRKAG3"
                     /note="AMPKG3"
                     /codon_start=1
                     /product="AMP-activated protein kinase gamma subunit"
                     /protein_id="AAF73987.1"
                     /db_xref="GI:8215682"
                     /translation="MSPLEQENSSSWPSPAVTSSSERIRGKRRAKALRWTRQKSVEEG
                     EPPGQGEGPRSRPTAESTGLEATFPKTTPLAQADPAGVGTPPTGWDCLPSDCTASAAG
                     SSTDDVELATEFPATEAWECELEGLLEERPALCLSPQAPFPKLGWDDELRKPGAQIYM
                     RFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEQHKIETWREIYLQGCFKPLVSISPNDSLFEAVYTLI
                     KNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIFGSLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNECGQVVGLYSRFDVIHLAAQQTYNHLDMSV
                     GEALRQRTLCLEGVLSCQPHESLGEVIDRIAREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT      460 a    622 c    562 g    471 t
ORIGIN


  Query Match           100.0%;  Score 101;  DB 9;  Length 2115;
  Best Local Similarity  100.0%;  Pred. No. 1e-20;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 gctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcac 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    906 GCTCAAGTTCCTGCACATCTTTGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCAC 965

Qy     61 tatccaagatttgggcatcggcacattccgagacttggctg 101
          |||||||||||||||||||||||||||||||||||||||||
Db    966 TATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCTG 1006


RESULT   5
HSA249977
LOCUS       HSA249977               2290 bp    mRNA    linear   PRI 07-APR-2000
DEFINITION  Homo sapiens mRNA for AMP-activated protein kinase gamma 3 subunit
            (AMPK gamma 3 gene).
ACCESSION   AJ249977
VERSION     AJ249977.1  GI:6688200
KEYWORDS    AMP-activated protein kinase; AMPK gamma 3 gene; gamma 3 subunit.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 2290)
  AUTHORS   Cheung,P.C., Salt,I.P., Davies,S.P., Hardie,D.G. and Carling,D.
  TITLE     Characterization of AMP-activated protein kinase gamma-subunit
            isoforms and their role in AMP binding
  JOURNAL   Biochem. J. 346 Pt 3, 659-669 (2000)
  MEDLINE   20164049
REFERENCE   2  (bases 1 to 2290)
  AUTHORS   Carling,D.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-OCT-1999) Carling D., Cellular Stress Group, MRC
            Clinical Sciences Centre, Hammersmith Hospital, DuCane Road,
            London, W12 0NN, UNITED KINGDOM
FEATURES             Location/Qualifiers
     source          1. .2290
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     gene            22. .1500
                     /gene="AMPK gamma 3"
     CDS             22. .1500
                     /gene="AMPK gamma 3"
                     /function="AMP-activated protein kinase regulatory
                     subunit"
                     /codon_start=1
                     /evidence=experimental
                     /product="AMP-activated protein kinase gamma 3 subunit"
                     /protein_id="CAB65117.1"
                     /db_xref="GI:6688201"
                     /translation="MEPGLEHALRRTPSWSSLGGSEHQEMSFLEQENSSSWPSPAVTS
                     SSERIRGKRRAKALRWTRQKSVEEGEPPGQGEGPRSRPAAESTGLEATFPKTTPLAQA
                     DPAGVGTPPTGWDCLPSDCTASAAGSSTDDVELATEFPATEAWECELEGLLEERPALC
                     LSPQAPFPKLGWDDELRKPGAQIYMRFIEEHTCYDAMATSSKLVIFDTMLEIKKAFFA
                     LVANGVRAAPLWDSKKQSFVGMLTITDFILVLHRYYRSPLVQIYEIEQHKIETWREIY
                     LQGCFKPLVSISPNDSLFEAVYTLIKNRIHRLPVLDPVSGNVLHILTHKRLLKFLHIF
                     GSLLPRPSFLYRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSALPVVNECGQV
                     VGLYSRFDVIHLAAQQTYNHLDMSVGEALRKRTLCLEGVLSCQPHESLGEVIDRIARE
                     QVHRLVLVDETQHLLGVVSLSDILQALVLSPAGIDPSGPEKI"
BASE COUNT      501 a    674 c    617 g    498 t
ORIGIN


  Query Match           100.0%;  Score 101;  DB 9;  Length 2290;
  Best Local Similarity  100.0%;  Pred. No. 1e-20;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      1 gctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcac 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1002 GCTCAAGTTCCTGCACATCTTTGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCAC 1061

Qy     61 tatccaagatttgggcatcggcacattccgagacttggctg 101
          |||||||||||||||||||||||||||||||||||||||||
Db   1062 TATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCTG 1102


RESULT   6
AX099774
LOCUS       AX099774                1867 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 1 from Patent WO0120003.
ACCESSION   AX099774
VERSION     AX099774.1  GI:13538808
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 1867)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 1 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
```

```
FEATURES             Location/Qualifiers
     source          1. .1867
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
     CDS             472. .1389
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35798.1"
                     /db_xref="GI:13538809"
                     /translation="MHFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAP
                     LWDSKKQSFVGMLTITDFILVLHRYYRSPLVQIYEIEEHKIETWREIYLQGCFKPLVS
                     ISPNDSLFEAVYALIKNRIHRLPVLDPVSGAVLHILTHKRLLKFLHIFGTLLPRPSFL
                     YRTIQDLGIGTFRDLAVVLETAPILTALDIFVDRRVSALPVVNETGQVVGLYSRFDVI
                     HLAAQQTYNHLDMNVGEALRQRTLCLEGVLSCQPHETLGEVIDRIVREQVHRLVLVDE
                     TQHLLGVVSLSDILQALVLSPAGIDALGA"
BASE COUNT      380 a    583 c    529 g    375 t
ORIGIN


  Query Match           92.7%;  Score 93.6;  DB 6;  Length 1867;
  Best Local Similarity   96.0%;  Pred. No. 1.9e-18;
  Matches   96;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
          ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Db    901 CTCAAGTTCCTGCACATCTTTGGCACCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACC 960

Qy     62 atccaagatttgggcatcggcacattccgagacttggctg 101
          |||||||||||||||||||||||||||||||||||| |
Db    961 ATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCCG 1000


RESULT   7
AF214520
LOCUS       AF214520                1873 bp    mRNA    linear   MAM 03-JUN-2000
DEFINITION  Sus scrofa AMP-activated protein kinase gamma subunit (PRKAG3)
            mRNA, complete cds.
ACCESSION   AF214520
VERSION     AF214520.1  GI:8215683
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 1873)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A., Thelander,M.,
            Rogel-Gaillard,C., Paul,S., Iannuccelli,N., Rask,L., Ronne,H.,
            Lundstrom,K., Reinsch,N., Gellin,J., Kalm,E., Roy,P.L., Chardon,P.
            and Andersson,L.
  TITLE     A mutation in PRKAG3 associated with excess glycogen content in pig
            skeletal muscle
  JOURNAL   Science 288 (5469), 1248-1251 (2000)
  MEDLINE   20280150
   PUBMED   10818001
REFERENCE   2  (bases 1 to 1873)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A.,
            Rogel-Gaillard,C., Paul,S., Gellin,J., Lundstrom,K., Reinsch,N.,
            Kalm,E., Le Roy,P., Chardon,P. and Andersson,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-1999) Dept Animal Breeding and Genetics, Swedish
            University of Agricultural Sciences, BMC box 597, Uppsala 751 24,
            Sweden
FEATURES             Location/Qualifiers
     source          1. .1873
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /chromosome="15"
                     /map="15q"
                     /tissue_type="skeletal muscle"
     gene            1. .1873
                     /gene="PRKAG3"
     CDS             1. .1395
                     /gene="PRKAG3"
                     /note="AMPKG3"
                     /codon_start=1
                     /product="AMP-activated protein kinase gamma subunit"
                     /protein_id="AAF73988.1"
                     /db_xref="GI:8215684"
                     /translation="MSFLEQGESRSWPSRAVTTSSERSHGDQGNKASRWTRQEDVEEG
                     GPPGPREGPQSRPVAESTGQEATFPKATPLAQAAPLAEVDNPPTERDILPSDCAASAS
                     DSNTDHLDLGIEFSASAASGDELGLVEEKPAPCPSPEVLLPRLGWDDELQKPGAQVYM
                     HFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEEHKIETWREIYLQGCFKPLVSISPNDSLFEAVYALI
                     KNRIHRLPVLDPVSGAVLHILTHKRLLKFLHIFGTLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNETGQVVGLYSRFDVIHLAAQQTYNHLDMNV
                     GEALRQRTLCLEGVLSCQPHETLGEVIDRIVREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT      382 a    580 c    535 g    376 t
ORIGIN


  Query Match           92.7%;  Score 93.6;  DB 4;  Length 1873;
  Best Local Similarity   96.0%;  Pred. No. 1.9e-18;
  Matches   96;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
          ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Db    907 CTCAAGTTCCTGCACATCTTTGGCACCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACC 966

Qy     62 atccaagatttgggcatcggcacattccgagacttggctg 101
          |||||||||||||||||||||||||||||||||||| |
Db    967 ATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCCG 1006


RESULT   8
AX099800
LOCUS       AX099800                1873 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 27 from Patent WO0120003.
ACCESSION   AX099800
VERSION     AX099800.1  GI:13538834
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 1873)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 27 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .1873
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
     CDS             1. .1395
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC35800.1"
                     /db_xref="GI:13538835"
                     /translation="MSFLEQGESRSWPSRAVTTSSERSHGDQGNKASRWTRQEDVEEG
                     GPPGPREGPQSRPVAESTGQEATFPKATPLAQAAPLAEVDNPPTERDILPSDCAASAS
                     DSNTDHLDLGIEFSASAASGDELGLVEEKPAPCPSPEVLLPRLGWDDELQKPGAQVYM
                     HFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEEHKIETWREIYLQGCFKPLVSISPNDSLFEAVYALI
                     KNRIHRLPVLDPVSGAVLHILTHKRLLKFLHIFGTLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNETGQVVGLYSRFDVIHLAAQQTYNHLDMNV
                     GEALRQRTLCLEGVLSCQPHETLGEVIDRIVREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT      382 a    580 c    535 g    376 t
ORIGIN
```

```
  Query Match             92.7%;  Score 93.6;  DB 6;  Length 1873;
  Best Local Similarity   96.0%;  Pred. No. 1.9e-18;
  Matches   96;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 ctcaagttcctgcacatctttggttccctgctgccccggcccctccttcctctaccgcact 61
          ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db    907 CTCAAGTTCCTGCACATCTTTGGCACCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACC 966

Qy     62 atccaagatttgggcatcggcacattccgagacttggctg 101
          |||||||||||||||||||||||||||||||||||||| |
Db    967 ATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCCG 1006


RESULT  9
AX099804
LOCUS       AX099804                2022 bp    DNA     linear   PAT 02-APR-2001
DEFINITION  Sequence 31 from Patent WO0120003.
ACCESSION   AX099804
VERSION     AX099804.1  GI:13538838
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 2022)
  AUTHORS   Andersson,L., Looft,C., Kalm,E., Milan,D., Robic,A.,
            Rogel-Gaillard,C., Iannuccelli,N., Gellin,J., le Roy,P. and
            Chardon,P.
  TITLE     Variants of the gamma chain of ampk, dna sequences encoding the
            same, and uses thereof
  JOURNAL   Patent: WO 0120003-A 31 22-MAR-2001;
            INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE (INRA) (FR) ;
            Andersson, Leif (SE) ; Looft, Christian (DE) ; Kalm, Ernst (DE)
FEATURES             Location/Qualifiers
     source          1. .2022
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
BASE COUNT      412 a    623 c    593 g    394 t
ORIGIN


  Query Match             92.7%;  Score 93.6;  DB 6;  Length 2022;
  Best Local Similarity   96.0%;  Pred. No. 1.9e-18;
  Matches   96;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy      2 ctcaagttcctgcacatctttggttccctgctgccccggcccctccttcctctaccgcact 61
          ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db   1057 CTCAAGTTCCTGCACATCTTTGGCACCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACC 1116

Qy     62 atccaagatttgggcatcggcacattccgagacttggctg 101
          |||||||||||||||||||||||||||||||||||||| |
Db   1117 ATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCCG 1156


RESULT  10
AX281580
LOCUS       AX281580                1722 bp    DNA     linear   PAT 03-NOV-2001
DEFINITION  Sequence 3 from Patent WO0177305.
ACCESSION   AX281580
VERSION     AX281580.1  GI:16608831
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (sites)
  AUTHORS   Andersson,L., Luthman,H. and Marklund,S.
  TITLE     Variants of the human amp-activated protein kinase gamma 3 subunit
  JOURNAL   Patent: WO 0177305-A 3 18-OCT-2001;
            Arexis AB (SE)
FEATURES             Location/Qualifiers
     source          1. .1722
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      321 a    504 c    534 g    363 t
ORIGIN


  Query Match             78.4%;  Score 79.2;  DB 6;  Length 1722;
  Best Local Similarity   91.3%;  Pred. No. 4.9e-14;
  Matches   84;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy     10 cctgcacatctttggttccctgctgccccggccctccttcctctaccgcactatccaaga 69
          ||  |  | |  ||||||||||||||||||||||||||||||||||||||||||||||||
Db   1510 CCATCCTAACCAGGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACTATCCAAGA 1569

Qy     70 tttgggcatcggcacattccgagacttggctg 101
          ||||||||||||||||||||||||||||||||
Db   1570 TTTGGGCATCGGCACATTCCGAGACTTGGCTG 1601


RESULT  11
AC027416/c
LOCUS       AC027416              152129 bp    DNA     linear   HTG 07-JUN-2000
DEFINITION  Homo sapiens clone RP11-504G11, WORKING DRAFT SEQUENCE, 32
            unordered pieces.
ACCESSION   AC027416
VERSION     AC027416.2  GI:8317289
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 152129)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-504G11
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 152129)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Jun 7, 2000 this sequence version replaced gi:7342115.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L7458
                Center clone name: 504_G_11
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
```

```
    Chemistry: Dye-terminator Big Dye; 100% of reads
    Assembly program: Phrap; version 0.960731
    Consensus quality: 135376 bases at least Q40
    Consensus quality: 143264 bases at least Q30
    Consensus quality: 146503 bases at least Q20
    Insert size: 161000; agarose-fp
    Insert size: 149029; sum-of-contigs
    Quality coverage: 3.1 in Q20 bases; agarose-fp
    Quality coverage: 3.3 in Q20 bases; sum-of-contigs
--------------.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 32 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*       1     1005: contig of 1005 bp in length
*    1006 1105: gap of      100 bp
*    1106     2402: contig of 1297 bp in length
*    2403 2502: gap of      100 bp
*    2503     3823: contig of 1321 bp in length
*    3824 3923: gap of      100 bp
*    3924     5020: contig of 1097 bp in length
*    5021 5120: gap of      100 bp
*    5121     6161: contig of 1041 bp in length
*    6162 6261: gap of      100 bp
*    6262     7547: contig of 1286 bp in length
*    7548 7647: gap of      100 bp
*    7648     9983: contig of 2336 bp in length
*    9984 10083: gap of      100 bp
*   10084    12556: contig of 2473 bp in length
*   12557 12656: gap of      100 bp
*   12657    15043: contig of 2387 bp in length
*   15044 15143: gap of      100 bp
*   15144    17123: contig of 1980 bp in length
*   17124 17223: gap of      100 bp
*   17224    19466: contig of 2243 bp in length
*   19467 19566: gap of      100 bp
*   19567    21928: contig of 2362 bp in length
*   21929 22028: gap of      100 bp
*   22029    24319: contig of 2291 bp in length
*   24320 24419: gap of      100 bp
*   24420    27059: contig of 2640 bp in length
*   27060 27159: gap of      100 bp
*   27160    30170: contig of 3011 bp in length
*   30171 30270: gap of      100 bp
*   30271    33968: contig of 3698 bp in length
*   33969 34068: gap of      100 bp
*   34069    38179: contig of 4111 bp in length
*   38180 38279: gap of      100 bp
*   38280    42366: contig of 4087 bp in length
*   42367 42466: gap of      100 bp
*   42467    46365: contig of 3899 bp in length
*   46366 46465: gap of      100 bp
*   46466    51285: contig of 4820 bp in length
*   51286 51385: gap of      100 bp
*   51386    55871: contig of 4486 bp in length
*   55872 55971: gap of      100 bp
*   55972    60595: contig of 4624 bp in length
*   60596 60695: gap of      100 bp
*   60696    66595: contig of 5900 bp in length
*   66596 66695: gap of      100 bp
*   66696    73218: contig of 6523 bp in length
*   73219 73318: gap of      100 bp
*   73319    77115: contig of 3797 bp in length
*   77116 77215: gap of      100 bp
*   77216    85022: contig of 7807 bp in length
*   85023 85122: gap of      100 bp
*   85123    93314: contig of 8192 bp in length
*   93315 93414: gap of      100 bp
*   93415   101193: contig of 7779 bp in length
*  101194 101293: gap of      100 bp
*  101294   113090: contig of 11797 bp in length
*  113091 113190: gap of      100 bp
*  113191   123496: contig of 10306 bp in length
*  123497 123596: gap of      100 bp
*  123597   137837: contig of 14241 bp in length
*  137838 137937: gap of      100 bp
*  137938   152129: contig of 14192 bp in length.
FEATURES             Location/Qualifiers
     source          1. .152129
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="RP11-504G11"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1005
                     /note="assembly_fragment"
     misc_feature    1106. .2402
                     /note="assembly_fragment"
     misc_feature    2503. .3823
                     /note="assembly_fragment"
     misc_feature    3924. .5020
                     /note="assembly_fragment"
     misc_feature    5121. .6161
                     /note="assembly_fragment"
     misc_feature    6262. .7547
                     /note="assembly_fragment"
     misc_feature    7648. .9983
                     /note="assembly_fragment"
     misc_feature    10084. .12556
                     /note="assembly_fragment"
     misc_feature    12657. .15043
                     /note="assembly_fragment"
     misc_feature    15144. .17123
                     /note="assembly_fragment"
     misc_feature    17224. .19466
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    19567. .21928
                     /note="assembly_fragment"
     misc_feature    22029. .24319
                     /note="assembly_fragment"
     misc_feature    24420. .27059
                     /note="assembly_fragment"
     misc_feature    27160. .30170
                     /note="assembly_fragment"
     misc_feature    30271. .33968
                     /note="assembly_fragment"
     misc_feature    34069. .38179
                     /note="assembly_fragment"
     misc_feature    38280. .42366
                     /note="assembly_fragment"
     misc_feature    42467. .46365
                     /note="assembly_fragment"
     misc_feature    46466. .51285
                     /note="assembly_fragment"
     misc_feature    51386. .55871
                     /note="assembly_fragment"
     misc_feature    55972. .60595
                     /note="assembly_fragment"
     misc_feature    60696. .66595
                     /note="assembly_fragment"
     misc_feature    66696. .73218
                     /note="assembly_fragment"
     misc_feature    73319. .77115
                     /note="assembly_fragment"
     misc_feature    77216. .85022
                     /note="assembly_fragment"
     misc_feature    85123. .93314
                     /note="assembly_fragment"
     misc_feature    93415. .101193
                     /note="assembly_fragment"
     misc_feature    101294. .113090
```

```
                     /note="assembly_fragment"

  Query Match             78.4%;  Score 79.2;  DB 2;  Length 152129;
  Best Local Similarity   91.3%;  Pred. No. 5.1e-14;
  Matches   84;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      10 cctgcacatctttggttccctgctgccccggccctccttcctctaccgcactatccaaga 69
           ||  |  | |   ||||||||||||||||||||||||||||||||||||||||||||
Db   35109 CCATCCTAACCAGGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACTATCCAAGA 35050

Qy      70 tttgggcatcggcacattccgagacttggctg 101
           |||||||||||||||||||||||||||||||
Db   35049 TTTGGGCATCGGCACATTCCGAGACTTGGCTG 35018


RESULT  12
AC073128/c
LOCUS       AC073128              196554 bp    DNA     linear   HTG 21-FEB-2001
DEFINITION  Homo sapiens chromosome 2 clone RP11-647O5, WORKING DRAFT SEQUENCE,
            17 unordered pieces.
ACCESSION   AC073128
VERSION     AC073128.3  GI:13027579
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 196554)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 196554)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-JUN-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Feb 21, 2001 this sequence version replaced gi:8469048.

            ----------------- Genome Center -----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: H_NH0647O05
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 98%
            Sequencing vector: plasmid; 0%
            Chemistry: Dye-primer ET; 98% of reads
            Chemistry: Dye-terminator Big Dye; 0% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 187795 bases at least Q40
            Consensus quality: 190513 bases at least Q30
            Consensus quality: 192099 bases at least Q20
            Insert size: 200000; agarose-fp
            Insert size: 194954; sum-of-contigs
            Quality coverage: 5.58 in Q20 bases; agarose-fp
            Quality coverage: 5.67 in Q20 bases; sum-of-contigs
            ----------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 17 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1157: contig of 1157 bp in length
            *     1158     1257: gap of unknown length
            *     1258     3600: contig of 2343 bp in length
            *     3601     3700: gap of unknown length
            *     3701     5103: contig of 1403 bp in length
            *     5104     5203: gap of unknown length
            *     5204     8524: contig of 3321 bp in length
            *     8525     8624: gap of unknown length
            *     8625    11856: contig of 3232 bp in length
            *    11857    11956: gap of unknown length
            *    11957    15783: contig of 3827 bp in length
            *    15784    15883: gap of unknown length
            *    15884    21906: contig of 6023 bp in length
            *    21907    22006: gap of unknown length
            *    22007    28887: contig of 6881 bp in length
            *    28888    28987: gap of unknown length
            *    28988    35255: contig of 6268 bp in length
            *    35256    35355: gap of unknown length
            *    35356    44642: contig of 9287 bp in length
            *    44643    44742: gap of unknown length
            *    44743    58275: contig of 13533 bp in length
            *    58276    58375: gap of unknown length
            *    58376    73816: contig of 15441 bp in length
            *    73817    73916: gap of unknown length
            *    73917    92140: contig of 18224 bp in length
            *    92141    92240: gap of unknown length
            *    92241   113337: contig of 21097 bp in length
            *   113338   113437: gap of unknown length
            *   113438   130325: contig of 16888 bp in length
            *   130326   130425: gap of unknown length
            *   130426   149287: contig of 18862 bp in length
            *   149288   149387: gap of unknown length
            *   149388   196554: contig of 47167 bp in length.
FEATURES             Location/Qualifiers
     source          1. .196554
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /clone="RP11-647O5"
     misc_feature    1. .1157
                     /note="assembly_name:Contig17"
     misc_feature    1258. .3600
                     /note="assembly_name:Contig18"
     misc_feature    3701. .5103
                     /note="assembly_name:Contig19
                     clone_end:T7
                     vector_side:right"
     misc_feature    5204. .8524
                     /note="assembly_name:Contig20"
     misc_feature    8625. .11856
                     /note="assembly_name:Contig21"
     misc_feature    11957. .15783
                     /note="assembly_name:Contig22"
     misc_feature    15884. .21906
                     /note="assembly_name:Contig23"
     misc_feature    22007. .28887
                     /note="assembly_name:Contig24"
     misc_feature    28988. .35255
                     /note="assembly_name:Contig25"
     misc_feature    35356. .44642
                     /note="assembly_name:Contig26"
     misc_feature    44743. .58275
                     /note="assembly_name:Contig27"
     misc_feature    58376. .73816
                     /note="assembly_name:Contig28"
     misc_feature    73917. .92140
                     /note="assembly_name:Contig29"
     misc_feature    92241. .113337
                     /note="assembly_name:Contig30"
     misc_feature    113438. .130325
                     /note="assembly_name:Contig31"
     misc_feature    130426. .149287
                     /note="assembly_name:Contig32"
     misc_feature    149388. .196554
                     /note="assembly_name:Contig33
                     clone_end:SP6
                     vector_side:right"
BASE COUNT    52296 a  46993 c  45889 g  49770 t   1606 others
ORIGIN
```

```
  Query Match             78.4%;  Score 79.2;  DB 2;  Length 196554;
  Best Local Similarity   91.3%;  Pred. No. 5.2e-14;
  Matches   84;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      10 cctgcacatctttggttccctgctgccccggccctccttcctctaccgcactatccaaga 69
           ||  |  | |   ||||||||||||||||||||||||||||||||||||||||||||||
Db   59610 CCATCCTAACCAGGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACTATCCAAGA 59551

Qy      70 tttgggcatcggcacattccgagacttggctg 101
           |||||||||||||||||||||||||||||||
Db   59550 TTTGGGCATCGGCACATTCCGAGACTTGGCTG 59519


RESULT  13
AC009974/c
LOCUS       AC009974               206854 bp    DNA     linear   PRI 09-JAN-2002
DEFINITION  Homo sapiens BAC clone RP11-459I19 from 2, complete sequence.
ACCESSION   AC009974
VERSION     AC009974.9  GI:16799058
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 206854)
  AUTHORS   Sulston,J.E. and Waterston,R.
  TITLE     Toward a complete human genome sequence
  JOURNAL   Genome Res. 8 (11), 1097-1108 (1998)
  MEDLINE   99063792
REFERENCE   2  (bases 1 to 206854)
  AUTHORS   Harris,A. and Cotton,M.
  TITLE     The sequence of Homo sapiens BAC clone RP11-459I19
  JOURNAL   Unpublished (2001)
REFERENCE   3  (bases 1 to 206854)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-SEP-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   4  (bases 1 to 206854)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-NOV-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   5  (bases 1 to 206854)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (03-JAN-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE   6  (bases 1 to 206854)
  AUTHORS   Waterston,R.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-JAN-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT     On Nov 8, 2001 this sequence version replaced gi:13431203.
            -------------- Genome Center
                Center: Washington University Genome Sequencing Center
                Center code: WUGSC
                Web site: http://genome.wustl.edu/gsc
                Contact: sapiens@watson.wustl.edu
            -------------- Summary Statistics
                Center project name: H_NH0459I19
            --------------.

            NOTICE:  This sequence may not represent the entire insert of this
            clone.  It may be shorter because we only sequence overlapping
            clone sections once, or longer because we provide a small overlap
            between neighboring data submissions.

            This sequence was finished as follows unless otherwise noted:
            all regions were double stranded, sequenced with an alternate
            chemistry, or covered by high quality data (i.e., phred quality >=
            30); an attempt was made to resolve all sequencing problems, such
            as compressions and repeats; all regions were covered by sequence
            from more than one subclone; and the assembly was confirmed by
            restriction digest.

            MAPPING INFORMATION:
            Mapping information for this clone was provided by Dr. John D.
            McPherson, Department of Genetics, Washington University, St. Louis
            MO.  For additional information about the map position of this
            sequence, see http://genome.wustl.edu/gsc

            SOURCE INFORMATION:
            The RPCI-11 human BAC library was made from the blood of one male
            donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frengen,E.,
            Tateno,M., Catanese,J.J. and de Jong,P.J. (1998) An improved
            approach for construction of bacterial artificial chromosome
            libraries.  Genomics 51:1-8.  The clone may be obtained either from
            Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
            and coworkers at the Roswell Park Cancer Institute
            (http://bacpac.med.buffalo.edu)
            VECTOR:  pBACe3.6
            NEIGHBORING SEQUENCE INFORMATION:
            The clone sequenced to the left is RP11-1077K22; the clone
            sequenced to the right is RP11-647O5.  Actual start of this clone
            is at base position 1 of RP11-459I19; actual end is at base
            position 206854 of RP11-459I19.

            Data from AC079810 and AC073128 was used to finish this clone,
            AC009974.  Polymorphisms have been identified between AC073128 and
            AC009974.  A single plasmid region exists beteen 38812-38903.  Am
            unresolved tandem in the HERV SVA exists between 184390-185163.
            PCR suggests that approximately 1700 bps are missing.
FEATURES             Location/Qualifiers
     source          1. .206854
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="2"
                     /map="2"
                     /clone="RP11-459I19"
                     /clone_lib="RPCI-11"
     misc_feature    1. .37
                     /note="match to EST BE314060 (NID:g9134719)"
     misc_feature    1. .37
                     /note="match to EST BI059713 (NID:g14467240)"
     misc_feature    1. .37
                     /note="match to EST BF183086 (NID:g11061273)"
     misc_feature    1. .37
                     /note="match to EST AL567345 (NID:g12920610)"
     misc_feature    1. .37
                     /note="match to EST AW880850 (NID:g8042860)"
     misc_feature    1. .37
                     /note="match to EST BF304755 (NID:g11251653)"
     misc_feature    1. .37
                     /note="similar to Homo sapiens EST BI114348
                     (NID:g14565249)"
     misc_feature    1. .37
                     /note="match to EST BG477625 (NID:g13409904)"
     misc_feature    1. .37
                     /note="match to EST BE047599 (NID:g8364652) tz39c01.y1"
     misc_feature    3. .37
                     /note="match to EST BE908408 (NID:g10402954)"
     misc_feature    4. .37
                     /note="match to EST AI670836 (NID:g4850567) wa04g10.x1"
     misc_feature    164. .662
                     /note="match to EST C05773 (NID:g1502549)"
     misc_feature    220. .221
                     /note="match to EST BG470047 (NID:g13402322)"
     misc_feature    281. .344
                     /note="match to EST BE047599 (NID:g8364652) tz39c01.y1"
     misc_feature    281. .344
```

```
                     /note="match to EST BE908408 (NID:g10402954)"
     misc_feature    281. .344
                     /note="match to EST BF304755 (NID:g11251653)"
     misc_feature    281. .344
                     /note="match to EST BF183086 (NID:g11061273)"
     misc_feature    281. .344
                     /note="similar to Homo sapiens EST BI114348
                     (NID:g14565249)"
     misc_feature    281. .344
                     /note="match to EST BG477625 (NID:g13409904)"
     misc_feature    281. .344
                     /note="match to EST BI059713 (NID:g14467240)"
     misc_feature    281. .344
                     /note="match to EST AL567345 (NID:g12920610)"
     misc_feature    281. .344
                     /note="match to EST BE314060 (NID:g9134719)"
     misc_feature    281. .344
                     /note="match to EST AI670836 (NID:g4850567) wa04g10.x1"
     misc_feature    281. .344
                     /note="match to EST AW880850 (NID:g8042860)"
     misc_feature    294. .344
                     /note="match to EST BG470047 (NID:g13402322)"
     misc_feature    complement(557. .558)
                     /note="match to EST AA043371 (NID:g1521226) zk53e10.r1"
     misc_feature    588. .929
                     /note="match to EST T64073 (NID:g667938) yc05d12.r1"
     misc_feature    594. .763
                     /note="match to EST BF304755 (NID:g11251653)"
     misc_feature    594. .763
                     /note="match to EST BG470047 (NID:g13402322)"
     misc_feature    594. .763
                     /note="match to EST AW880850 (NID:g8042860)"
     misc_feature    594. .731
                     /note="match to EST BE314060 (NID:g9134719)"
     misc_feature    594. .764
                     /note="match to EST BI059713 (NID:g14467240)"
     misc_feature    594. .763
                     /note="match to EST BG477625 (NID:g13409904)"
     misc_feature    594. .742
                     /note="match to EST BE047599 (NID:g8364652) tz39c01.y1"
     misc_feature    594. .763
                     /note="match to EST BE908408 (NID:g10402954)"
     misc_feature    594. .763
                     /note="match to EST AA043371 (NID:g1521226) zk53e10.r1"
     misc_feature    594. .763
                     /note="match to EST AI670836 (NID:g4850567) wa04g10.x1"
     misc_feature    594. .763
                     /note="match to EST BF183086 (NID:g11061273)"
     misc_feature    594. .763
                     /note="similar to Homo sapiens EST BI114348
                     (NID:g14565249)"
     misc_feature    599. .763
                     /note="match to EST AL567345 (NID:g12920610)"
     misc_feature    617. .1084
                     /note="match to EST AA481361 (NID:g2210913) zv44e01.r1"
     misc_feature    622. .763
                     /note="match to EST AI860958 (NID:g5514574) wl56f05.x1"
     misc_feature    622. .763
                     /note="similar to Mus musculus EST AI196847 (NID:g3749453)
                     ui67e01.x1"
     misc_feature    684. .763
                     /note="match to EST BG992568 (NID:g14396638)"
     misc_feature    962. .1084
                     /note="match to EST AI656812 (NID:g4740791) tt54b06.x1"
     misc_feature    967. .1084
                     /note="match to EST BE908408 (NID:g10402954)"
     misc_feature    967. .1085
                     /note="match to EST BF304755 (NID:g11251653)"
     misc_feature    967. .1091
                     /note="match to EST AA043371 (NID:g1521226) zk53e10.r1"
     misc_feature    967. .1071

Query Match          78.4%;  Score 79.2;  DB 9;  Length 206854;
  Best Local Similarity   91.3%;  Pred. No. 5.2e-14;
  Matches   84;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      10 cctgcacatctttggttccctgctgccccggccctccttcctctaccgcactatccaaga 69
           ||  |  | |   ||||||||||||||||||||||||||||||||||||||||||||||
Db 165190 CCATCCTAACCAGGGTTCCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACTATCCAAGA 165131

Qy      70 tttgggcatcggcacattccgagacttggctg 101
           |||||||||||||||||||||||||||||||
Db 165130 TTTGGGCATCGGCACATTCCGAGACTTGGCTG 165099


RESULT  14
AF214521
LOCUS       AF214521                5888 bp    DNA     linear   MAM 03-JUN-2000
DEFINITION  Sus scrofa AMPK gamma subunit (PRKAG3) gene, complete cds.
ACCESSION   AF214521
VERSION     AF214521.1  GI:8215685
KEYWORDS    .
SOURCE      pig.
  ORGANISM  Sus scrofa
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE   1  (bases 1 to 5888)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A., Thelander,M.,
            Rogel-Gaillard,C., Paul,S., Iannuccelli,N., Rask,L., Ronne,H.,
            Lundstrom,K., Reinsch,N., Gellin,J., Kalm,E., Roy,P.L., Chardon,P.
            and Andersson,L.
  TITLE     A mutation in PRKAG3 associated with excess glycogen content in pig
            skeletal muscle
  JOURNAL   Science 288 (5469), 1248-1251 (2000)
  MEDLINE   20280150
   PUBMED   10818001
REFERENCE   2  (bases 1 to 5888)
  AUTHORS   Milan,D., Jeon,J.T., Looft,C., Amarger,V., Robic,A.,
            Rogel-Gaillard,C., Paul,S., Gellin,J., Lundstrom,K., Reinsch,N.,
            Kalm,E., Le Roy,P., Chardon,P. and Andersson,L.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-DEC-1999) Dept Animal Breeding and Genetics, Swedish
            University of Agricultural Sciences, BMC Box 597, Uppsala 751 24,
            Sweden
FEATURES             Location/Qualifiers
     source          1. .5888
                     /organism="Sus scrofa"
                     /db_xref="taxon:9823"
                     /chromosome="15"
                     /map="15q"
     mRNA            join(<1. .154,515. .918,1809. .1890,2349. .2407,2509. .2554,
                     2771. .2825,3027. .3153,3286. .3451,4578. .4615,4791. .4937,
                     5294. .>5410)
                     /gene="PRKAG3"
                     /product="AMPK gamma subunit"
     gene            <1. .>5410
                     /gene="PRKAG3"
     CDS             join(1. .154,515. .918,1809. .1890,2349. .2407,2509. .2554,
                     2771. .2825,3027. .3153,3286. .3451,4578. .4615,4791. .4937,
                     5294. .5410)
                     /gene="PRKAG3"
                     /codon_start=1
                     /product="AMPK gamma subunit"
                     /protein_id="AAF73989.1"
                     /db_xref="GI:8215686"
                     /translation="MSFLEQGESRSWPSRAVTTSSERSHGDQGNKASRWTRQEDVEEG
                     GPPGPREGPQSRPVAESTGQEATFPKATPLAQAAPLAEVDNPPTERDILPSDCAASAS
                     DSNTDHLDLGIEFSASAASGDELGLVEEKPAPCPSPEVLLPRLGWDDELQKPGAQVYM
                     HFMQEHTCYDAMATSSKLVIFDTMLEIKKAFFALVANGVRAAPLWDSKKQSFVGMLTI
                     TDFILVLHRYYRSPLVQIYEIEEHKIETWREIYLQGCFKPLVSISPNDSLFEAVYALI
                     KNRIHRLPVLDPVSGAVLHILTHKRLLKFLHIFGTLLPRPSFLYRTIQDLGIGTFRDL
                     AVVLETAPILTALDIFVDRRVSALPVVNETGQVVGLYSRFDVIHLAAQQTYNHLDMNV
                     GEALRQRTLCLEGVLSCQPHETLGEVIDRIVREQVHRLVLVDETQHLLGVVSLSDILQ
                     ALVLSPAGIDALGA"
BASE COUNT     1274 a   1634 c   1638 g   1339 t      3 others
ORIGIN
```

```
  Query Match            75.2%;  Score 76;  DB 4;  Length 5888;
  Best Local Similarity   85.0%;  Pred. No. 4.7e-13;
  Matches   85;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy      2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
          | |  || ||  | || |   ||  ||||||||||||||||||||||||||||||||||
Db   3265 CACCTGTCCCAACCCAACCAGGGCACCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACC 3324

Qy     62 atccaagatttgggcatcggcacattccgagacttggctg 101
          |||||||||||||||||||||||||||||||||||||||| |
Db   3325 ATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCCG 3364


RESULT  15
AF336381/c
LOCUS       AF336381               227724 bp    DNA     linear   HTG 02-APR-2001
DEFINITION  Mus musculus chromosome 1 clone PAC510; PAC457, *** SEQUENCING IN
            PROGRESS ***, 3 unordered pieces.
ACCESSION   AF336381
VERSION     AF336381.1  GI:13507298
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 227724)
  AUTHORS   Rump,A., Hayes,C., Brown,S.D.M. and Rosenthal,A.
  TITLE     Mouse chromosome 1 genomic sequence
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 227724)
  AUTHORS   Rump,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (17-JAN-2001) Genome Analysis, Institute of Molecular
            Biotechnology, Beutenberstr. 11, Jena 07745, Germany
COMMENT     * NOTE: This is a 'working draft' sequence. It currently
            * consists of 3 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1    17869: contig of 17869 bp in length
            *    17870    17969: gap of unknown length
            *    17970    32746: contig of 14777 bp in length
            *    32747    32846: gap of unknown length
            *    32847   227724: contig of 194878 bp in length.
FEATURES             Location/Qualifiers
     source          1. .227724
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /chromosome="1"
                     /clone="PAC510; PAC457"
BASE COUNT    57663 a  55423 c  56238 g  58199 t    201 others
ORIGIN


  Query Match            70.3%;  Score 71;  DB 2;  Length 227724;
  Best Local Similarity   93.7%;  Pred. No. 1.7e-11;
  Matches   74;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy     23 ggttccctgctgccccggccctccttcctctaccgcactatccaagatttgggcatcggc 82
          ||| ||||| ||||||||||||||||||||| |||||||||||||| ||||||||||||
Db   5531 GGTGCCCTGTTGCCCCGGCCCTCCTTCCTCTGCCGCACTATCCAAGACTTGGGCATCGGC 5472

Qy     83 acattccgagacttggctg 101
          |||||||||| |||||||
Db   5471 ACATTCCGAGATTTGGCTG 5453
```

```
Search completed: October  3, 2002, 14:50:23
Job time: 12143 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 16:30:46 ; Search time 719.93 Seconds
(without alignments)
240.868 Million cell updates/sec

Title: US-09-826-581-5_COPY_1000_1100
Perfect score: 101
Sequence: 1 gctcaagttcctgcacatct.....cacattccgagacttggctg 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_032802:*
1: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	101	100.0	547	22	ABA08485	Human AMP-activate
	2	101	100.0	1647	22	AAH43685	PRKAG3 cDNA. Homo
	3	101	100.0	2109	22	AAD03296	Human AMPK gamma s
	4	101	100.0	2115	22	AAD03320	Human AMPK gamma s
	5	93.6	92.7	1867	22	AAD03295	Pig AMPK gamma sub
	6	93.6	92.7	1873	22	AAD03319	Pig AMPK gamma sub
	7	93.6	92.7	2022	22	AAD03321	Sus scrofa PRKAG3
	8	79.2	78.4	1722	22	AAH43683	PRKAG3 intron 4 -
	9	30.8	30.5	1435	20	AAX06882	Disease associated
	10	30.8	30.5	1467	23	AAS84265	DNA encoding novel
	11	30.8	30.5	2223	22	AAH14839	Human cDNA sequenc
c	12	29.6	29.3	566	22	AAS05559	Mammalian vestibul
	13	29.6	29.3	1326	21	AAC37238	Arabidopsis thalia
c	14	29.6	29.3	2289	23	ABL08199	Drosophila melanog
c	15	29.6	29.3	4363	23	ABL08198	Drosophila melanog
	16	29.4	29.1	200	24	AAS17161	Mouse melanocyte h
	17	29.4	29.1	1259	14	AAQ51226	Mouse MSH-R gene.
	18	29.4	29.1	1260	19	AAV19135	Mouse melanocyte s
	19	29.4	29.1	1260	20	AAV63702	Mouse melanocortin
	20	29.4	29.1	1260	24	AAS17159	Mouse melanocyte h
	21	29.4	29.1	5007	23	ABL07647	Drosophila melanog
c	22	29.4	29.1	12403	23	ABL07646	Drosophila melanog
	23	29	28.7	4413	23	ABL02825	Drosophila melanog
	24	29	28.7	6765	23	ABL02824	Drosophila melanog
	25	29	28.7	7461	23	ABL04284	Drosophila melanog
	26	28.8	28.5	398	22	AAF64424	Novel human polynu
	27	28.8	28.5	9979	22	AAK81194	Human immune/haema
	28	28.8	28.5	13673	22	AAK76548	Human immune/haema
	29	28.8	28.5	13673	22	AAK81195	Human immune/haema
	30	28.4	28.1	1682	21	AAC98901	Human pancreatic c
c	31	28.2	27.9	635	22	AAL02279	Human reproductive
c	32	28.2	27.9	2378	23	ABL06405	Drosophila melanog
	33	28.2	27.9	2530	21	AAC39424	Arabidopsis thalia
c	34	28.2	27.9	2588	23	ABL21171	Drosophila melanog
c	35	28.2	27.9	2857	22	AAH15945	Human cDNA sequenc
c	36	28.2	27.9	3238	23	AAS77333	DNA encoding novel
c	37	28.2	27.9	4378	23	ABL06404	Drosophila melanog
c	38	28.2	27.9	4732	23	ABL21170	Drosophila melanog
c	39	28	27.7	301	23	AAS58263	cDNA #939 encoding
	40	28	27.7	737	22	AAH03932	Human cDNA clone (
	41	28	27.7	3125	22	AAH15387	Human cDNA sequenc
	42	28	27.7	3156	19	AAV18471	T-cell surface ant
	43	28	27.7	3156	20	AAZ27969	Human CD97 protein
	44	28	27.7	3156	21	AAA95996	Human CD97 nucleot
c	45	28	27.7	12043	24	AAS95361	Human intercellula

ALIGNMENTS

RESULT 1
ABA08485
ID ABA08485 standard; cDNA; 547 BP.
XX
AC ABA08485;
XX
DT 11-JAN-2002 (first entry)
XX
DE Human AMP-activated protein kinase subunit homologue cDNA, SEQ ID NO:261.
XX
KW Human; cytokine; cell proliferation; cell differentiation; growth factor;
KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
KW chronic inflammatory condition; proliferative retinopathy;
KW atherosclerosis; coronary heart disease; arterial ischaemia;
KW bone disorder; osteoporosis; vascular growth disorder;
KW tissue regeneration; wound healing; infection; immune disorder;
KW cell culture; drug screening; gene therapy; antiinflammatory;
KW antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
KW cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;
KW antifungal; vulnerary; antiulcer; ss.
XX
OS Homo sapiens.
XX
PN WO200157188-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US03800.
XX

```
PR   03-FEB-2000; 2000US-0496914.
PR   27-APR-2000; 2000US-0560875.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT,  Liu C,  Drmanac RT;
XX
DR   WPI; 2001-457740/49.
DR   P-PSDB; ABB11241.
XX
PT   Human proteins and DNA encoding sequences useful for preventing,
PT   treating or ameliorating a medical condition in a mammalian subject
PT   e.g. arthritis and cancer -
XX
PS   Claim 1; Page 429; 1963pp; English.
XX
CC   Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
CC   sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
CC   invention also relates to vectors and recombinant host cells comprising a
CC   nucleotide of the invention, methods of producing the novel polypeptides,
CC   antibodies against the polypeptides, methods of detecting the nucleotides
CC   or polypeptides in a sample, and methods of identifying compounds which
CC   bind to polypeptides of the invention. Although novel, many of the
CC   polypeptides of the invention have homology to known proteins, thereby
CC   giving an insight into their probable biological activities, and hence
CC   potential therapeutic applications. The polypeptides of the invention may
CC   have various activities, including cytokine, cell proliferation or cell
CC   differentiation activities; stem cell growth factor activity;
CC   haematopoiesis regulatory activity; tissue growth activity;
CC   immunomodulatory activity; activin- or inhibin-related activities;
CC   chemotactic or chemokinetic activities; haemostatic, thrombotic or
CC   thrombolytic activities; receptor or ligand activities; or may be
CC   involved in oncogenesis, cancer cell proliferation or metastasis.
CC   Depending on their biological activities, polypeptides and nucleotides of
CC   the invention are useful for preventing, treating or ameliorating medical
CC   conditions, e.g., by protein or gene therapy. Such conditions include
CC   cancers, haematopoietic disorders (e.g.,myeloid or lymphoid cell
CC   disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
CC   proliferative retinopathy, atherosclerosis, coronary heart disease,
CC   arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal
CC   vascular growth. Polypeptides involved with tissue regeneration and
CC   repair (or nucleic acids encoding them) may be used to promote wound
CC   healing (e.g., of burns, incisions and ulcers), while those with
CC   immunomodulatory activities may be used in the treatment of viral,
CC   bacterial and fungal infections in addition to immune disorders.
CC   Polypeptides with growth factor activity may be used in cell cultures to
CC   promote cell growth. For example, such polypeptides may be used to
CC   manipulate stem cells in culture to give rise to neuroepithelial cells
CC   that can be used to augment or replace cells damaged by illness,
CC   autoimmune disease or accidental damage. The polypeptides and nucleotides
CC   may also be used in the diagnosis of the above conditions, and in drug
CC   screening techniques. The present sequence represents a cDNA encoding a
CC   novel human polypeptide of the invention.
XX
SQ   Sequence 547 BP; 112 A; 172 C; 133 G; 130 T; 0 other;

  Query Match             100.0%;  Score 101;  DB 22;  Length 547;
  Best Local Similarity   100.0%;  Pred. No. 7.9e-23;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcac 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     263 gctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcac 322

Qy      61 tatccaagatttgggcatcggcacattccgagacttggctg 101
           |||||||||||||||||||||||||||||||||||||||||
Db     323 tatccaagatttgggcatcggcacattccgagacttggctg 363


RESULT   2
AAH43685
ID   AAH43685 standard; cDNA; 1647 BP.
XX
AC   AAH43685;
XX
DT   21-JAN-2002  (first entry)
XX
DE   PRKAG3 cDNA.
XX
KW   Human; AMP-activated protein kinase gamma 3 subunit; PRKAG3; variant;
KW   metabolic disease; diabetes; obesity; substitution; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             20..1489
FT                   /*tag= a
FT                   /product= "PRKAG3"
FT   variation       230
FT                   /*tag= b
FT                   /label= "C230G"
FT                   /note= "Causes P71A"
FT   variation       559
FT                   /*tag= c
FT                   /label= "T559C"
FT                   /note= "Silent variation"
FT   variation       1037
FT                   /*tag= d
FT                   /label= "C1037T"
FT                   /note= "Causes R340W"
XX
PN   WO200177305-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-SE00765.
XX
PR   07-APR-2000; 2000US-195665P.
XX
PA   (AREX-) AREXIS AB.
XX
PI   Andersson L,  Luthman H,  Marklund S;
XX
DR   WPI; 2001-657170/75.
DR   P-PSDB; QQB47679.
XX
PT   New variants of human AMP-activated protein kinase gamma3 subunit
PT   associated with a metabolic disease e.g. diabetes or obesity and method
PT   for determining a risk estimate of diseases in subject by detecting the
PT   variant -
XX
PS   Disclosure; Fig 5; 25pp; English.
XX
CC   This sequence represents the full length cDNA encoding the human
CC   AMP-activated protein kinase gamma 3 subunit (PRKAG3). Detecting
CC   the presence of the PRKAG3 DNA, or a variant, is useful in determining
CC   a risk estimate of a metabolic disease, such as diabetes or obesity,
CC   in a subject.  The variation may occur in exons 3, 4 or 10.  In exon
CC   3 variation may be a substitution of a G for a C at nucleotide 320,
CC   resulting in the amino acid substitution P71A; in exon 4 variation may
CC   be a substitution of a T for a C at nucleotide 550; and in exon 10
CC   variation may be a substitution of a T for a C at nucleotide 1037,
CC   resulting in the amino acid substitution R340W.  There may also be
CC   nucleotide variation in intron 6.  The numbering of these
CC   variations is based on the full length cDNA as given, rather than on
CC   position 1 of the open reading frame.
XX
SQ   Sequence 1647 BP; 346 A; 502 C; 462 G; 337 T; 0 other;

  Query Match             100.0%;  Score 101;  DB 22;  Length 1647;
  Best Local Similarity   100.0%;  Pred. No. 1e-22;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcac 60
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1000 gctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcac 1059

Qy      61 tatccaagatttgggcatcggcacattccgagacttggctg 101
           |||||||||||||||||||||||||||||||||||||||||
Db    1060 tatccaagatttgggcatcggcacattccgagacttggctg 1100


RESULT   3
AAD03296
ID   AAD03296 standard; DNA; 2109 BP.
XX
AC   AAD03296;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Human AMPK gamma subunit muscle-specific isoform, PRKAG3 cDNA.
XX
KW   Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW   PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW   genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW   cystathione beta synthase; CBS; cardiant; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1..471
FT                   /*tag= a
FT   CDS             472..1389
FT                   /*tag= b
FT                   /product= "Human Prkag3 protein"
FT   3'UTR           1390..2109
FT                   /*tag= c
XX
PN   WO200120003-A2.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-EP09896.
XX
PR   10-SEP-1999;   99EP-0402236.
PR   18-MAY-2000; 2000EP-0401388.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA   (ANDE/) ANDERSSON L.
PA   (LOOF/) LOOFT C.
PA   (KALM/) KALM E.
XX
PI   Andersson L,  Looft C,  Kalm E,  Milan D,  Robic A,  Rogel-Gaillard C;
PI   Iannuccelli N,  Gellin J,  Le Roy P,  Chardon P;
XX
DR   WPI; 2001-244810/25.
DR   P-PSDB; AAE00221.
XX
PT   New variants of the gamma subunit of vertebrate adenosine
PT   monophosphate-activated kinase for diagnosis or treatment of disorders
PT   associated with energy metabolism such as diabetes, obesity, and
PT   myopathy  -
XX
PS   Claim 12; Fig 2; 71pp; English.
XX
CC   The present sequence is a cDNA encoding human adenosine monophosphate
CC   (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC   PRKAG3. Mutation in Prkag3 results in an altered regulation of
CC   carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is
CC   useful as therapeutic for treating carbohydrate metabolism disorders such
CC   as diabetes, obesity, and disorders associated with muscle metabolism
CC   such as myopathy and cardiovascular diseases, to modulate AMPK
CC   activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC   and its functionally altered mutants are useful for the diagnostic
CC   evaluation, genetic testing and prognosis of a metabolic disorder,
CC   preferably a carbohydrate metabolism disorder. Primers that can detect
CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
CC   from the expression of a functionally altered allele of PRKAG3.
CC   Transgenic animal and host cell transformed with PRKAG3 or a
CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC   screening compounds able to modulate AMPK activity. Nucleic acid
CC   encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
CC   of PRKAG3 and is useful in gene therapy.
XX
SQ   Sequence 2109 BP; 458 A; 621 C; 560 G; 470 T; 0 other;


  Query Match             100.0%;  Score 101;  DB 22;  Length 2109;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-22;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcac 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     900 gctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcac 959

Qy      61 tatccaagatttgggcatcggcacattccgagacttggctg 101
           |||||||||||||||||||||||||||||||||||||||||
Db     960 tatccaagatttgggcatcggcacattccgagacttggctg 1000


RESULT   4
AAD03320
ID   AAD03320 standard; cDNA; 2115 BP.
XX
AC   AAD03320;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Human AMPK gamma subunit muscle-specific isoform, complete PRKAG3 cDNA.
XX
KW   Human; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW   PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW   genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW   cystathione beta synthase; CBS; cardiant; gene therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1395
FT                   /*tag= a
FT                   /product= "Human complete Prkag3 protein"
XX
PN   WO200120003-A2.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-EP09896.
XX
PR   10-SEP-1999;   99EP-0402236.
PR   18-MAY-2000; 2000EP-0401388.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA   (ANDE/) ANDERSSON L.
PA   (LOOF/) LOOFT C.
PA   (KALM/) KALM E.
XX
PI   Andersson L,  Looft C,  Kalm E,  Milan D,  Robic A,  Rogel-Gaillard C;
PI   Iannuccelli N,  Gellin J,  Le Roy P,  Chardon P;
XX
DR   WPI; 2001-244810/25.
DR   P-PSDB; AAE00223.
XX
PT   New variants of the gamma subunit of vertebrate adenosine
PT   monophosphate-activated kinase for diagnosis or treatment of disorders
PT   associated with energy metabolism such as diabetes, obesity, and
PT   myopathy  -
XX
PS   Claim 12; Page 65-68; 71pp; English.
```

```
XX
CC   The present sequence is a cDNA encoding human adenosine monophosphate
CC   (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC   complete PRKAG3. Mutation in Prkag3 results in an altered regulation of
CC   carbohydrate metabolism, particularly in skeletal muscle. PRKAG3 is
CC   useful as therapeutic for treating carbohydrate metabolism disorders such
CC   as diabetes, obesity, and disorders associated with muscle metabolism
CC   such as myopathy and cardiovascular diseases, to modulate AMPK
CC   activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC   and its functionally altered mutants are useful for the diagnostic
CC   evaluation, genetic testing and prognosis of a metabolic disorder,
CC   preferably a carbohydrate metabolism disorder. Primers that can detect
CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
CC   from the expression of a functionally altered allele of PRKAG3.
CC   Transgenic animal and host cell transformed with PRKAG3 or a
CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC   screening compounds able to modulate AMPK activity. Nucleic acid
CC   encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
CC   of PRKAG3 and is useful in gene therapy.
XX
SQ   Sequence 2115 BP; 460 A; 622 C; 562 G; 471 T; 0 other;


  Query Match             100.0%;  Score 101;  DB 22;  Length 2115;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-22;
  Matches  101;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy       1 gctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcac 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     906 gctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcac 965

Qy      61 tatccaagatttgggcatcggcacattccgagacttggctg 101
           |||||||||||||||||||||||||||||||||||||||||
Db     966 tatccaagatttgggcatcggcacattccgagacttggctg 1006


RESULT   5
AAD03295
ID   AAD03295 standard; cDNA; 1867 BP.
XX
AC   AAD03295;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Pig AMPK gamma subunit muscle-specific isoform, PRKAG3 cDNA.
XX
KW   Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW   PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW   genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW   cystathione beta synthase; CBS; cardiant; gene therapy; RN locus;
KW   chromosome 15; ss.
XX
OS   Sus scrofa.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1..471
FT                   /*tag= a
FT   CDS             472..1389
FT                   /*tag= b
FT                   /product= "Sus scrofa PRKAG3 protein"
FT   3'UTR           1390..1867
FT                   /*tag= c
XX
PN   WO200120003-A2.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-EP09896.
XX
PR   10-SEP-1999;   99EP-0402236.
PR   18-MAY-2000; 2000EP-0401388.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA   (ANDE/) ANDERSSON L.
PA   (LOOF/) LOOFT C.
PA   (KALM/) KALM E.
XX
PI   Andersson L,  Looft C,  Kalm E,  Milan D,  Robic A,  Rogel-Gaillard C;
PI   Iannuccelli N,  Gellin J,  Le Roy P,  Chardon P;
XX
DR   WPI; 2001-244810/25.
DR   P-PSDB; AAE00220.
XX
PT   New variants of the gamma subunit of vertebrate adenosine
PT   monophosphate-activated kinase for diagnosis or treatment of disorders
PT   associated with energy metabolism such as diabetes, obesity, and
PT   myopathy  -
XX
PS   Claim 12; Fig 2; 71pp; English.
XX
CC   The present sequence is a cDNA encoding pig adenosine monophosphate
CC   (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC   PRKAG3. Prkag3 gene is located in the RN locus of chromosome 15.
CC   Mutation in Prkag3 results in an altered regulation of carbohydrate
CC   metabolism, particularly in skeletal muscle. PRKAG3 is useful as
CC   therapeutic for treating carbohydrate metabolism disorders such as
CC   diabetes, obesity, and disorders associated with muscle metabolism
CC   such as myopathy and cardiovascular diseases, to modulate AMPK
CC   activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC   and its functionally altered mutants are useful for the diagnostic
CC   evaluation, genetic testing and prognosis of a metabolic disorder,
CC   preferably a carbohydrate metabolism disorder. Primers that can detect
CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
CC   from the expression of a functionally altered allele of PRKAG3.
CC   Transgenic animal and host cell transformed with PRKAG3 or a
CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC   screening compounds able to modulate AMPK activity. Nucleic acid
CC   encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
CC   of PRKAG3 and is useful in gene therapy.
XX
SQ   Sequence 1867 BP; 380 A; 583 C; 529 G; 375 T; 0 other;


  Query Match             92.7%;  Score 93.6;  DB 22;  Length 1867;
  Best Local Similarity   96.0%;  Pred. No. 2.4e-20;
  Matches   96;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           |||||||||||||||||||||||  ||||||||||||||||||||||||||||||||||
Db     901 ctcaagttcctgcacatctttggcaccctgctgccccggccctccttcctctaccgcacc 960

Qy      62 atccaagatttgggcatcggcacattccgagacttggctg 101
           |||||||||||||||||||||||||||||||||||||| |
Db     961 atccaagatttgggcatcggcacattccgagacttggccg 1000


RESULT   6
AAD03319
ID   AAD03319 standard; cDNA; 1873 BP.
XX
AC   AAD03319;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Pig AMPK gamma subunit muscle-specific isoform, complete PRKAG3 cDNA.
XX
KW   Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW   PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW   genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW   cystathione beta synthase; CBS; cardiant; gene therapy; RN locus;
KW   chromosome 15; ss.
XX
```

```
OS   Sus scrofa.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1395
FT                   /*tag= a
FT                   /product= "Sus scrofa complete Prkag3 protein"
XX
PN   WO200120003-A2.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-EP09896.
XX
PR   10-SEP-1999;   99EP-0402236.
PR   18-MAY-2000; 2000EP-0401388.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA   (ANDE/) ANDERSSON L.
PA   (LOOF/) LOOFT C.
PA   (KALM/) KALM E.
XX
PI   Andersson L,  Looft C,  Kalm E,  Milan D,  Robic A,  Rogel-Gaillard C;
PI   Iannuccelli N,  Gellin J,  Le Roy P,  Chardon P;
XX
DR   WPI; 2001-244810/25.
DR   P-PSDB; AAE00222.
XX
PT   New variants of the gamma subunit of vertebrate adenosine
PT   monophosphate-activated kinase for diagnosis or treatment of disorders
PT   associated with energy metabolism such as diabetes, obesity, and
PT   myopathy  -
XX
PS   Claim 12; Page 62-64; 71pp; English.
XX
CC   The present sequence is a cDNA encoding pig adenosine monophosphate
CC   (AMP)-activated kinase (AMPK) gamma subunit muscle-specific isoform,
CC   complete PRKAG3. Prkag3 gene is located in the RN locus of chromosome
CC   15. Mutation in Prkag3 results in an altered regulation of carbohydrate
CC   metabolism, particularly in skeletal muscle. PRKAG3 is useful as
CC   therapeutic for treating carbohydrate metabolism disorders such as
CC   diabetes, obesity, and disorders associated with muscle metabolism
CC   such as myopathy and cardiovascular diseases, to modulate AMPK
CC   activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC   and its functionally altered mutants are useful for the diagnostic
CC   evaluation, genetic testing and prognosis of a metabolic disorder,
CC   preferably a carbohydrate metabolism disorder. Primers that can detect
CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
CC   from the expression of a functionally altered allele of PRKAG3.
CC   Transgenic animal and host cell transformed with PRKAG3 or a
CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC   screening compounds able to modulate AMPK activity. Nucleic acid
CC   encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
CC   of PRKAG3 and is useful in gene therapy.
XX
SQ   Sequence 1873 BP; 382 A; 580 C; 535 G; 376 T; 0 other;


  Query Match             92.7%;  Score 93.6;  DB 22;  Length 1873;
  Best Local Similarity   96.0%;  Pred. No. 2.4e-20;
  Matches   96;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           |||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||
Db     907 ctcaagttcctgcacatctttggcaccctgctgccccggccctccttcctctaccgcacc 966

Qy      62 atccaagatttgggcatcggcacattccgagacttggctg 101
           |||||||||||||||||||||||||||||||||||||| |
Db     967 atccaagatttgggcatcggcacattccgagacttggccg 1006


RESULT   7
```

```
AAD03321
ID   AAD03321 standard; DNA; 2022 BP.
XX
AC   AAD03321;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Sus scrofa PRKAG3 splice variant DNA.
XX
KW   Pig; gamma subunit; adenosine monophosphate-activated kinase; AMPK;
KW   PRKAG3; diabetes; obesity; myopathy; cardiovascular disease; anorectic;
KW   genetic testing; carbohydrate metabolism disorder; skeletal muscle;
KW   cystathione beta synthase; CBS; cardiant; gene therapy; ds.
XX
OS   Sus scrofa.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..1545
FT                   /*tag= a
FT                   /product= "Sus scrofa Prkag3 splice variant"
XX
PN   WO200120003-A2.
XX
PD   22-MAR-2001.
XX
PF   11-SEP-2000; 2000WO-EP09896.
XX
PR   10-SEP-1999;   99EP-0402236.
PR   18-MAY-2000; 2000EP-0401388.
XX
PA   (INRG ) INRA INST NAT RECH AGRONOMIQUE.
PA   (ANDE/) ANDERSSON L.
PA   (LOOF/) LOOFT C.
PA   (KALM/) KALM E.
XX
PI   Andersson L,  Looft C,  Kalm E,  Milan D,  Robic A,  Rogel-Gaillard C;
PI   Iannuccelli N,  Gellin J,  Le Roy P,  Chardon P;
XX
DR   WPI; 2001-244810/25.
DR   P-PSDB; AAE00224.
XX
PT   New variants of the gamma subunit of vertebrate adenosine
PT   monophosphate-activated kinase for diagnosis or treatment of disorders
PT   associated with energy metabolism such as diabetes, obesity, and
PT   myopathy  -
XX
PS   Claim 12; Page 69; 71pp; English.
XX
CC   The present sequence is pig adenosine monophosphate (AMP)-activated
CC   kinase (AMPK) gamma subunit muscle-specific isoform, PRKAG3 splice
CC   variant DNA. Prkag3 gene is located in the RN locus of chromosome 15.
CC   Mutation in Prkag3 results in an altered regulation of carbohydrate
CC   metabolism, particularly in skeletal muscle. PRKAG3 is useful as
CC   therapeutic for treating carbohydrate metabolism disorders such as
CC   diabetes, obesity, and disorders associated with muscle metabolism
CC   such as myopathy and cardiovascular diseases, to modulate AMPK
CC   activity, and for restoring a normal AMPK function. PRKAG3 sequence
CC   and its functionally altered mutants are useful for the diagnostic
CC   evaluation, genetic testing and prognosis of a metabolic disorder,
CC   preferably a carbohydrate metabolism disorder. Primers that can detect
CC   a genetic polymorphic marker linked to a sequence encoding PRKAG3, are
CC   useful for detecting a dysfunction of carbohydrate metabolism resulting
CC   from the expression of a functionally altered allele of PRKAG3.
CC   Transgenic animal and host cell transformed with PRKAG3 or a
CC   heterotrimeric AMPK consisting of PRKAG3 or its mutant, are useful for
CC   screening compounds able to modulate AMPK activity. Nucleic acid
CC   encoding PRKAG3 is useful for detecting mutations in a Prkag3 gene, or
CC   in a sequence encoding the first cystathione beta synthase (CBS) domain
CC   of PRKAG3 and is useful in gene therapy.
XX
SQ   Sequence 2022 BP; 412 A; 623 C; 593 G; 394 T; 0 other;
```

```
  Query Match             92.7%;  Score 93.6;  DB 22;  Length 2022;
  Best Local Similarity   96.0%;  Pred. No. 2.4e-20;
  Matches   96;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           ||||||||||||||||||||||  ||||||||||||||||||||||||||||||||
Db    1057 ctcaagttcctgcacatctttggcaccctgctgccccggccctccttcctctaccgcacc 1116

Qy      62 atccaagatttgggcatcggcacattccgagacttggctg 101
           |||||||||||||||||||||||||||||||||||| |
Db    1117 atccaagatttgggcatcggcacattccgagacttggccg 1156


RESULT   8
AAH43683
ID   AAH43683 standard; DNA; 1722 BP.
XX
AC   AAH43683;
XX
DT   21-JAN-2002  (first entry)
XX
DE   PRKAG3 intron 4 - intron 10.
XX
KW   Human; AMP-activated protein kinase gamma 3 subunit; PRKAG3; variant;
KW   metabolic disease; diabetes; obesity; substitution; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   intron          1..13
FT                   /*tag= a
FT                   /number = "Intron 4"
FT                   /note= "3' portion of intron 4"
FT   exon            14..95
FT                   /*tag= b
FT                   /number= "Exon 5"
FT   intron          96..552
FT                   /*tag= c
FT                   /number= "Intron 5"
FT   exon            553..611
FT                   /*tag= d
FT                   /number= "Exon 6"
FT   intron          612..736
FT                   /*tag= e
FT                   /number= "Intron 6"
FT   exon            737..782
FT                   /*tag= f
FT                   /number= "Exon 7"
FT   intron          783..986
FT                   /*tag= g
FT                   /number= "Intron 7"
FT   exon            987..1041
FT                   /*tag= h
FT                   /number= "Exon 8"
FT   intron          1042..1242
FT                   /*tag= i
FT                   /number= "Intron 8"
FT   exon            1243..1369
FT                   /*tag= j
FT                   /number= "Exon 9"
FT   intron          1370..1522
FT                   /*tag= k
FT                   /number= "Intron 9"
FT   exon            1523..1688
FT                   /*tag= l
FT                   /number= "Exon 10"
FT   intron          1689..1722
FT                   /*tag= i
FT                   /number= "Intron 10"
FT                   /note= "5' portion of intron 10"
XX
PN   WO200177305-A2.
XX
PD   18-OCT-2001.
XX
PF   06-APR-2001; 2001WO-SE00765.
XX
PR   07-APR-2000; 2000US-195665P.
XX
PA   (AREX-) AREXIS AB.
XX
PI   Andersson L,  Luthman H,  Marklund S;
XX
DR   WPI; 2001-657170/75.
XX
PT   New variants of human AMP-activated protein kinase gamma3 subunit
PT   associated with a metabolic disease e.g. diabetes or obesity and method
PT   for determining a risk estimate of diseases in subject by detecting the
PT   variant -
XX
PS   Example 1; Fig 3; 25pp; English.
XX
CC   The sequences given in AAH43681-84 represents genomic fragments
CC   encoding the human AMP-activated protein kinase gamma 3 subunit
CC   (PRKAG3). Detecting the presence of the PRKAG3 DNA, or a variant,
CC   is useful in determining a risk estimate of a metabolic disease,
CC   such as diabetes or obesity, in a subject.  The variation may occur
CC   in exons 3, 4 or 10.  In exon 3 variation may be a substitution of
CC   a G for a C at nucleotide 320, resulting in the amino acid
CC   substitution P71A; in exon 4 variation may be a substitution of a
CC   T for a C at nucleotide 550; and in exon 10 variation may be a
CC   substitution of a T for a C at nucleotide 1037, resulting in the
CC   amino acid substitution R340W.  There may also be nucleotide variation
CC   in intron 6.
XX
SQ   Sequence 1722 BP; 321 A; 504 C; 534 G; 363 T; 0 other;


  Query Match             78.4%;  Score 79.2;  DB 22;  Length 1722;
  Best Local Similarity   91.3%;  Pred. No. 9e-16;
  Matches   84;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy      10 cctgcacatctttggttccctgctgccccggccctccttcctctaccgcactatccaaga 69
             ||  |  | |   ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1510 ccatcctaaccagggttccctgctgccccggccctccttcctctaccgcactatccaaga 1569

Qy      70 tttgggcatcggcacattccgagacttggctg 101
           |||||||||||||||||||||||||||||||
Db    1570 tttgggcatcggcacattccgagacttggctg 1601


RESULT   9
AAX06882
ID   AAX06882 standard; cDNA; 1435 BP.
XX
AC   AAX06882;
XX
DT   26-APR-1999  (first entry)
XX
DE   Disease associated protein kinase DAPK-7 cDNA.
XX
KW   DAPK-7; disease associated protein kinase; human; diagnosis;
KW   therapy; adult respiratory distress syndrome; allergy; asthma;
KW   arteriosclerosis; bronchitis; emphysema; hypereosinophilia;
KW   myocardial inflammation; pericardial inflammation; anaemia;,
KW   rheumatoid arthritis; Addison's disease; AIDS; atherosclerosis;
KW   atopic dermatitis; dermatomyositis; diabetes mellitus;
KW   glomerulonephritis; gout; Grave's disease; lupus erythematosus;
KW   multiple sclerosis; myasthenia gravis; osteoarthritis;
KW   osteoporosis; pancreatitis; polycystic kidney disease;
KW   polymyositis; scleroderma; Sjorgren's syndrome;
KW   autoimmune thyroiditis; cancer; infection; trauma;
KW   cell proliferation; ss.
XX
OS   Homo sapiens.
XX
```

```
FH   Key             Location/Qualifiers
FT   CDS             265..1503
FT                   /*tag= a
XX
PN   WO9858052-A2.
XX
PD   23-DEC-1998.
XX
PF   19-JUN-1998;    98WO-US12813.
XX
PR   19-JUN-1997;    97US-0878989.
XX
PA   (INCY-) INCYTE PHARM INC.
XX
PI   Bandman O, Corley NC, Goli SK, Guegler KJ, Hillman JL;
PI   Lal P, Shah P;
XX
DR   WPI; 1999-080952/07.
DR   P-PSDB; AAW88438.
XX
PT   New disease associated protein kinases - used to stimulate cell
PT   proliferation and to treat the immune response and cancer
XX
PS   Claim 5; Page 66-67; 93pp; English.
XX
CC   This cDNA sequence codes for human disease associated protein
CC   kinase DAPK-7 (see AAW88438). DAPK-7 cDNA was first identified in
CC   the PENITUT01 cDNA library using a computer search for amino acid
CC   alignments, and a consensus sequence was derived from the extended
CC   and overlapping Incyte clones 3075712/HEARNOT01, 842220/PROSTUT05,
CC   1364747/SCORNON02, 145972 and 145802/PENITUT01 and 1479332/CORPNOT02.
CC   DAPK-7 shows 73% homology with the human foetal liver AMPK gamma
CC   subunit (G1 1335856), and is associated with cDNA libraries which
CC   are immortalised or cancerous and show inflammatory or immune
CC   responses. The invention provides disease associated protein kinases
CC   DAPK-1 to DAPK-7 (see AAW88432-38) and cDNA clones encoding them (see
CC   AAX06831-36 and AAX06882), as well as expression vectors, host cells,
CC   agonists, antagonists and antibodies. The invention further provides
CC   uses of such products in the diagnosis, prevention and treatment of
CC   diseases associated with cell proliferation, especially cancer or an
CC   immune response (claimed). Conditions that may be treated include
CC   adult respiratory distress syndrome, allergies, asthma,
CC   arteriosclerosis, bronchitis, emphysema, hypereosinophilia,
CC   myocardial or pericardial inflammation, rheumatoid arthritis,
CC   Addison's disease, AIDS, anaemia, atherosclerosis, various diseases
CC   of the digestive system, atopic dermatitis, dermatomyositis, diabetes
CC   mellitus, glomerulonephritis, gout, Grave's disease, lupus
CC   erythematosus, multiple sclerosis, myasthenia gravis,
CC   osteoarthritis, osteoporosis, pancreatitis, polycystic kidney
CC   disease, polymyositis, scleroderma, Sjorgren's syndrome, autoimmune
CC   thyroiditis, complications of cancer, extracorporeal circulation,
CC   viral, bacterial, fungal, parasitic, protozoal and helminthic
CC   infections, and trauma (disclosed). The DAPK nucleic acids are
CC   also used in a method for detection of DAPK expression levels in a
CC   biological sample.
XX
SQ   Sequence 1435 BP; 421 A; 298 C; 331 G; 385 T; 0 other;

  Query Match             30.5%;  Score 30.8;  DB 20;  Length 1435;
  Best Local Similarity   57.1%;  Pred. No. 2.3;
  Matches   56;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           |||||||||||| ||  | |||   ||     ||||   |||  ||||| |  | |  |
Db     617 ctcaagttcctccagctttttatgtctgatatgccaaagcctgccttcatgaagcagaac 676

Qy      62 atccaagatttgggcatcggcacattccgagacttggc 99
             |  | ||  | || || || || | ||   || | ||
Db     677 ctggatgagcttggaataggaacgtaccacaacattgc 714

RESULT 10
AAS84265
ID   AAS84265 standard; cDNA; 1467 BP.
XX
AC   AAS84265;
XX
DT   13-FEB-2002  (first entry)
XX
DE   DNA encoding novel human diagnostic protein #20069.
XX
KW   Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW   food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS   Homo sapiens.
XX
PN   WO200175067-A2.
XX
PD   11-OCT-2001.
XX
PF   30-MAR-2001; 2001WO-US08631.
XX
PR   31-MAR-2000; 2000US-0540217.
PR   23-AUG-2000; 2000US-0649167.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Drmanac RT, Liu C, Tang YT;
XX
DR   WPI; 2001-639362/73.
DR   P-PSDB; ABG20078.
XX
PT   New isolated polynucleotide and encoded polypeptides, useful in
PT   diagnostics, forensics, gene mapping, identification of mutations
PT   responsible for genetic disorders or other traits and to assess
PT   biodiversity -
XX
PS   Claim 1; SEQ ID No 20069; 103pp; English.
XX
CC   The invention relates to isolated polynucleotide (I) and
CC   polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC   polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC   and gene mapping, and in recombinant production of (II). The
CC   polynucleotides are also used in diagnostics as expressed sequence tags
CC   for identifying expressed genes. (I) is useful in gene therapy techniques
CC   to restore normal activity of (II) or to treat disease states involving
CC   (II). (II) is useful for generating antibodies against it, detecting or
CC   quantitating a polypeptide in tissue, as molecular weight markers and as
CC   a food supplement. (II) and its binding partners are useful in medical
CC   imaging of sites expressing (II). (I) and (II) are useful for treating
CC   disorders involving aberrant protein expression or biological activity.
CC   The polypeptide and polynucleotide sequences have applications in
CC   diagnostics, forensics, gene mapping, identification of mutations
CC   responsible for genetic disorders or other traits to assess biodiversity
CC   and to produce other types of data and products dependent on DNA and
CC   amino acid sequences. AAS64197-AAS94564 represent novel human
CC   diagnostic coding sequences of the invention.
CC   Note: The sequence data for this patent did not appear in the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 1467 BP; 416 A; 327 C; 343 G; 381 T; 0 other;

  Query Match             30.5%;  Score 30.8;  DB 23;  Length 1467;
  Best Local Similarity   57.1%;  Pred. No. 2.3;
  Matches   56;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           |||||||||||| ||  | |||   ||     ||||   |||  ||||| |  | |  |
Db     715 ctcaagttcctccagctttttatgtctgatatgccaaagcctgccttcatgaagcagaac 774

Qy      62 atccaagatttgggcatcggcacattccgagacttggc 99
             |  | ||  | || || || || | ||   || | ||
Db     775 ctggatgagcttggaataggaacgtaccacaacattgc 812
```

```
RESULT  11
AAH14839
ID   AAH14839 standard; cDNA; 2223 BP.
XX
AC   AAH14839;
XX
DT   26-JUN-2001  (first entry)
XX
DE   Human cDNA sequence SEQ ID NO:12660.
XX
KW   Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   EP1074617-A2.
XX
PD   07-FEB-2001.
XX
PF   28-JUL-2000; 2000EP-0116126.
XX
PR   29-JUL-1999;   99JP-0248036.
PR   27-AUG-1999;   99JP-0300253.
PR   11-JAN-2000; 2000JP-0118776.
PR   02-MAY-2000; 2000JP-0183767.
PR   09-JUN-2000; 2000JP-0241899.
XX
PA   (HELI-) HELIX RES INST.
XX
PI   Ota T,  Isogai T,  Nishikawa T,  Hayashi K,  Saito K,  Yamamoto J;
PI   Ishii S,  Sugiyama T,  Wakamatsu A,  Nagai K,  Otsuki T;
XX
DR   WPI; 2001-318749/34.
XX
PT   Primer sets for synthesizing polynucleotides, particularly the 5602
PT   full-length cDNAs defined in the specification, and for the detection
PT   and/or diagnosis of the abnormality of the proteins encoded by the
PT   full-length cDNAs -
XX
PS   Claim 8; SEQ ID 12660; 2537pp + CD ROM; English.
XX
CC   The present invention describes primer sets for synthesising 5602
CC   full-length cDNAs defined in the specification. Where a primer set
CC   comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC   to the complementary strand of a polynucleotide which comprises one of
CC   the 5602 nucleotide sequences defined in the specification, where the
CC   oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC   of an oligonucleotide comprising a sequence complementary to the
CC   complementary strand of a polynucleotide which comprises a 5'-end
CC   sequence and an oligonucleotide comprising a sequence complementary to a
CC   polynucleotide which comprises a 3'-end sequence, where the
CC   oligonucleotide comprises at least 15 nucleotides and the combination of
CC   the 5'-end sequence/3'-end sequence is selected from those defined in
CC   the specification. The primer sets can be used in antisense therapy and
CC   in gene therapy. The primers are useful for synthesising polynucleotides,
CC   particularly full-length cDNAs. The primers are also useful for the
CC   detection and/or diagnosis of the abnormality of the proteins encoded by
CC   the full-length cDNAs. The primers allow obtaining of the full-length
CC   cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC   AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC   AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC   represent oligonucleotides, all of which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 2223 BP; 633 A; 431 C; 496 G; 663 T; 0 other;

  Query Match             30.5%;  Score 30.8;  DB 22;  Length 2223;
  Best Local Similarity   57.1%;  Pred. No. 2.5;
  Matches   56;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           ||||||||||| ||  | |||   ||     ||||   |||  ||||| |  | |  |
Db     649 ctcaagttcctccagctttttatgtctgatatgccaaagcctgccttcatgaagcagaac 708

Qy      62 atccaagatttgggcatcggcacattccgagacttggc 99
           |  | ||  | || || || || | ||   || | ||
Db     709 ctggatgagcttggaataggaacgtaccacaacattgc 746


RESULT  12
AAS05559/c
ID   AAS05559 standard; DNA; 566 BP.
XX
AC   AAS05559;
XX
DT   07-SEP-2001  (first entry)
XX
DE   Mammalian vestibular system geotactic behaviour modulator gene #159.
XX
KW   Mammalian vestibular system; invertebrate; geotactic behaviour; vertigo;
KW   graviperceptive disorder; motion sickness; labyrinthitis; syphilis; ds;
KW   Meniere's disease; acoustic neuroma; multiple sclerosis; epilepsy;
KW   trauma; infection of the middle ear; ototoxic agent exposure.
XX
OS   Drosophila melanogaster .
XX
PN   WO200140519-A2.
XX
PD   07-JUN-2001.
XX
PF   01-DEC-2000; 2000WO-US32639.
XX
PR   02-DEC-1999;   99US-0168579.
PR   26-SEP-2000; 2000US-0669751.
XX
PA   (NEUR-) NEUROSCIENCES RES FOUND INC.
XX
PI   Greenspan RJ;
XX
DR   WPI; 2001-356159/37.
XX
PT   New isolated nucleic acid having mammalian vestibular system-modulating
PT   activity useful in the treatment of disorders such as motion sickness
PT   and vertigo -
XX
PS   Claim 59; Page 144; 179pp; English.
XX
CC   The sequences shown in AAS05401-AAS05661 represent DNA with mammalian
CC   vestibular system-modulating activity. The DNA sequences can be used in a
CC   method whereby a first and second strain of an invertebrate is obtained,
CC   and both are subjected to conditions in which the strains exhibit
CC   different geotactic behaviour. Genes that are differentially expressed in
CC   the first strain relative to the second strain are then identified.
CC   Mammalian genes having substantially the same nucleic acid sequence as
CC   these modulate the mammalian vestibular system. Compounds containing
CC   these genes are used to decrease the symptoms of graviperceptive
CC   disorders such as motion sickness, vertigo, labyrinthitis, Meniere's
CC   disease, acoustic neuroma, multiple sclerosis, syphilis, trauma,
CC   infection of the middle ear, exposure to ototoxic agents and epilepsy.
XX
SQ   Sequence 566 BP; 191 A; 129 C; 161 G; 85 T; 0 other;

  Query Match             29.3%;  Score 29.6;  DB 22;  Length 566;
  Best Local Similarity   68.3%;  Pred. No. 4.4;
  Matches   41;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      13 gcacatctttggttccctgctgccccggccctccttcctctaccgcactatccaagattt 72
           || ||||   |||  |   |||  || || || ||||||||||| || |||||  ||| |
Db     544 GCTCATCAAAGGTGACTACCTGTTCCTGCTCTTCTTCCTCTACCTCATTATCCCTGATCT 485


RESULT  13
AAC37238
```

ID AAC37238 standard; DNA; 1326 BP.
XX
AC AAC37238;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 16667.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.

```
PR   23-AUG-1999;   99US-0149930.
PR   25-AUG-1999;   99US-0150566.
PR   26-AUG-1999;   99US-0150884.
PR   27-AUG-1999;   99US-0151065.
PR   27-AUG-1999;   99US-0151066.
PR   27-AUG-1999;   99US-0151080.
PR   30-AUG-1999;   99US-0151303.
PR   31-AUG-1999;   99US-0151438.
PR   01-SEP-1999;   99US-0151930.
PR   07-SEP-1999;   99US-0152363.
PR   10-SEP-1999;   99US-0153070.
PR   13-SEP-1999;   99US-0153758.
PR   15-SEP-1999;   99US-0154018.
PR   16-SEP-1999;   99US-0154039.
PR   20-SEP-1999;   99US-0154779.
PR   22-SEP-1999;   99US-0155139.
PR   23-SEP-1999;   99US-0155486.
PR   24-SEP-1999;   99US-0155659.
PR   28-SEP-1999;   99US-0156458.
PR   29-SEP-1999;   99US-0156596.
PR   04-OCT-1999;   99US-0157117.
PR   05-OCT-1999;   99US-0157753.
PR   06-OCT-1999;   99US-0157865.
PR   07-OCT-1999;   99US-0158029.
PR   08-OCT-1999;   99US-0158232.
PR   12-OCT-1999;   99US-0158369.
PR   13-OCT-1999;   99US-0159293.
PR   13-OCT-1999;   99US-0159294.
PR   13-OCT-1999;   99US-0159295.
PR   14-OCT-1999;   99US-0159329.
PR   14-OCT-1999;   99US-0159330.
PR   14-OCT-1999;   99US-0159331.
PR   14-OCT-1999;   99US-0159637.
PR   14-OCT-1999;   99US-0159638.
PR   18-OCT-1999;   99US-0159584.
PR   21-OCT-1999;   99US-0160741.
PR   21-OCT-1999;   99US-0160767.
PR   21-OCT-1999;   99US-0160768.
PR   21-OCT-1999;   99US-0160770.
PR   21-OCT-1999;   99US-0160814.
PR   21-OCT-1999;   99US-0160815.
PR   22-OCT-1999;   99US-0160980.
PR   22-OCT-1999;   99US-0160981.
PR   22-OCT-1999;   99US-0160989.
PR   25-OCT-1999;   99US-0161404.
PR   25-OCT-1999;   99US-0161405.
PR   25-OCT-1999;   99US-0161406.
PR   26-OCT-1999;   99US-0161359.
PR   26-OCT-1999;   99US-0161360.
PR   26-OCT-1999;   99US-0161361.
PR   28-OCT-1999;   99US-0161920.
PR   28-OCT-1999;   99US-0161992.
PR   28-OCT-1999;   99US-0161993.
PR   29-OCT-1999;   99US-0162142.

  Query Match             29.3%;  Score 29.6;  DB 21;  Length 1326;
  Best Local Similarity   64.7%;  Pred. No. 5.4;
  Matches   44;  Conservative    0;  Mismatches   24;  Indels    0;  Gaps    0;

Qy       8 ttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcactatccaa 67
           ||| ||| | |||| |  || |  |||||   ||||  ||  |||||   | |  |||||
Db     193 ttcttgctcttcttcgtctctcctctgccatcgcccagctctctctagttcgcgttccaa 252

Qy      68 gatttggg 75
           | || |||
Db     253 gtttaggg 260


RESULT  14
ABL08199/c
ID   ABL08199 standard; cDNA; 2289 BP.
XX
AC   ABL08199;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster expressed polynucleotide SEQ ID NO 19079.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ss.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
PD   27-SEP-2001.
XX
PF   23-MAR-2001; 2001WO-US09231.
XX
PR   23-MAR-2000; 2000US-191637P.
PR   11-JUL-2000; 2000US-0614150.
XX
PA   (PEKE ) PE CORP NY.
XX
PI   Venter JC,  Adams M,  Li PWD,  Myers EW;
XX
DR   WPI; 2001-656860/75.
DR   P-PSDB; ABB64096.
XX
PT   New isolated nucleic acid detection reagent for detecting 1000 or more
PT   genes from Drosophila and for elucidating cell signalling and cell-cell
PT   interactions -
XX
PS   Claim 1; SEQ ID NO 19079; 21pp + Sequence Listing; English.
XX
CC   The invention relates to an isolated nucleic acid detection reagent
CC   capable of detecting 1000 or more genes from Drosophila. The invention is
CC   useful in developmental biology and in elucidating cell signalling and
CC   cell-cell interactions in higher eukaryotes for the development of
CC   insecticides, therapeutics and pharmaceutical drugs. The invention
CC   discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC   sequences (ABL01840-ABL16175) and the encoded proteins
CC   (ABB57737-ABB72072).
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 2289 BP; 694 A; 541 C; 589 G; 465 T; 0 other;


  Query Match             29.3%;  Score 29.6;  DB 23;  Length 2289;
  Best Local Similarity   68.3%;  Pred. No. 6.1;
  Matches   41;  Conservative    0;  Mismatches   19;  Indels    0;  Gaps    0;

Qy      13 gcacatctttggttccctgctgccccggccctccttcctctaccgcactatccaagattt 72
           || ||||   |||  |   |||  || || || ||||||||||| || |||||  ||| |
Db     544 GCTCATCAAAGGTGACTACCTGTTCCTGCTCTTCTTCCTCTACCTCATTATCCCTGATCT 485


RESULT  15
ABL08198/c
ID   ABL08198 standard; cDNA; 4363 BP.
XX
AC   ABL08198;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Drosophila melanogaster expressed polynucleotide SEQ ID NO 19076.
XX
KW   Drosophila; developmental biology; cell signalling; insecticide;
KW   pharmaceutical; gene; ss.
XX
OS   Drosophila melanogaster.
XX
PN   WO200171042-A2.
XX
```

PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
DR P-PSDB; ABB64095.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 19076; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 4363 BP; 1461 A; 939 C; 974 G; 989 T; 0 other;

Query Match 29.3%; Score 29.6; DB 23; Length 4363;
Best Local Similarity 68.3%; Pred. No. 7;
Matches 41; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

```
Qy      13 gcacatctttggttccctgctgccccggccctccttcctctaccgcactatccaagattt 72
           || ||||   |||  |   |||  || || || ||||||||||| || |||||  ||| |
Db    1544 GCTCATCAAAGGTGACTACCTGTTCCTGCTCTTCTTCCTCTACCTCATTATCCCTGATCT 1485
```

Search completed: October 3, 2002, 16:30:48
Job time: 14318 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 16:15:08 ; Search time 5701.1 Seconds
(without alignments)
239.110 Million cell updates/sec

Title: US-09-826-581-5_COPY_1000_1100
Perfect score: 101
Sequence: 1 gctcaagttcctgcacatct.....cacattccgagacttggctg 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	93.6	92.7	572	10	BI344527	BI344527 373008 MA
	2	59	58.4	413	9	AA178898	AA178898 zp38d10.r
	3	48.8	48.3	633	10	BJ072114	BJ072114 BJ072114
	4	38.8	38.4	514	10	BG037921	BG037921 dc69g08.y
c	5	34.8	34.5	918	10	BG027175	BG027175 602295858
c	6	32.2	31.9	451	9	AW147353	AW147353 da01h11.y
c	7	32.2	31.9	470	10	BJ036397	BJ036397 BJ036397
c	8	32.2	31.9	580	9	AV914030	AV914030 AV914030
	9	31.4	31.1	684	10	BG101440	BG101440 uy72c12.y
	10	31.2	30.9	536	10	BM488662	BM488662 pgm2n.pk0
	11	31.2	30.9	595	10	BM487789	BM487789 pgm2n.pk0
	12	31.2	30.9	647	10	BM440762	BM440762 pgr1n.pk0
	13	31.2	30.9	649	9	AJ395115	AJ395115 AJ395115
	14	31.2	30.9	758	9	AJ396118	AJ396118 AJ396118
	15	30.8	30.5	413	10	H64260	H64260 yr70c12.r1
	16	30.8	30.5	451	10	W39604	W39604 zc19b04.r1
	17	30.8	30.5	474	10	BE803695	BE803695 sr62e03.y
	18	30.8	30.5	506	10	BI760420	BI760420 603045431
	19	30.8	30.5	525	10	BG148719	BG148719 uu86d02.y
	20	30.8	30.5	601	10	BE846246	BE846246 232075 BA
	21	30.8	30.5	764	9	AU138307	AU138307 AU138307
	22	30.8	30.5	849	10	BI657718	BI657718 603283104
	23	30.8	30.5	886	10	BG572824	BG572824 602594365
	24	30.8	30.5	909	9	AL558594	AL558594 AL558594
	25	30.8	30.5	951	9	AL513784	AL513784 AL513784
	26	30.8	30.5	967	9	AL524822	AL524822 AL524822
	27	30.4	30.1	312	10	C06172	C06172 C06172 Huma
c	28	30.2	29.9	924	12	CNS029MX	AL187458 Tetraodon
	29	30	29.7	663	10	BG415737	BG415737 HVSMEk000
	30	29.8	29.5	784	9	AU080009	AU080009 AU080009
	31	29.8	29.5	964	10	BG035134	BG035134 602324830
	32	29.6	29.3	322	9	AV525978	AV525978 AV525978
c	33	29.6	29.3	619	10	BE978092	BE978092 bs73g07.y
c	34	29.6	29.3	626	10	BF528264	BF528264 602043043
c	35	29.6	29.3	715	10	BF488250	BF488250 AT23475.5
c	36	29.6	29.3	754	9	AI063338	AI063338 GH03085.5
c	37	29.6	29.3	836	10	BI659310	BI659310 603301205
c	38	29.4	29.1	388	12	AQ535079	AQ535079 RPCI-11-3
	39	29.4	29.1	559	9	AI062936	AI062936 GH02321.5
	40	29.4	29.1	559	9	AI062937	AI062937 GH02322.5
c	41	29.4	29.1	588	10	BM340317	BM340317 MEST321-B
c	42	29.4	29.1	699	12	BH447226	BH447226 BOHNG10TF
	43	29.4	29.1	888	9	AL581895	AL581895 AL581895
c	44	29.2	28.9	224	9	AA045087	AA045087 zk77g11.r
	45	29.2	28.9	388	10	BI774231	BI774231 466266 MA

ALIGNMENTS

RESULT 1
BI344527
LOCUS BI344527 572 bp mRNA linear EST 30-JUL-2001
DEFINITION 373008 MARC 2PIG Sus scrofa cDNA 5', mRNA sequence.
ACCESSION BI344527
VERSION BI344527.1 GI:15037807
KEYWORDS EST.
SOURCE pig.
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 572)
AUTHORS Fahrenkrug,S.C., Freking,B.A., Rohrer,G.A., Smith,T.P.L., Casas,E., Stone,R.T., Heaton,M.P., Grosse,W.M., Bennett,G.A., Laegreid,W.W. and Keele,J.W.
TITLE Design and use of two pooled tissue normalized cDNA libraries for EST discovery in swine
JOURNAL Unpublished (2000)
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR PRimers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTTCCCAGTCACGACG
Plate: 119 row: I column: 11
Seq primer: ATTTAGGTGACACTATAG.
FEATURES Location/Qualifiers
source 1. .572
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone_lib="MARC 2PIG"
/tissue_type="pooled"
/lab_host="DH10B"
/note="Vector: pCMV SPORT6; Site_1: XbaI; Site_2: XhoI;

```
                    Library made from pooled tissue from testis, ovary,
                    endometrium, hypothalamus, pituitary, and placenta."
BASE COUNT       106 a    190 c    151 g    124 t      1 others
ORIGIN


  Query Match             92.7%;  Score 93.6;  DB 10;  Length 572;
  Best Local Similarity   96.0%;  Pred. No. 7.9e-16;
  Matches   96;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           ||||||||||||||||||||  |||||||||||||||||||||||||||||||||||||
Db     163 CTCAAGTTCCTGCACATCTTTGGCACCCTGCTGCCCCGGCCCTCCTTCCTCTACCGCACC 222

Qy      62 atccaagatttgggcatcggcacattccgagacttggctg 101
           |||||||||||||||||||||||||||||||||||| |
Db     223 ATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCCG 262


RESULT   2
AA178898
LOCUS       AA178898                413 bp    mRNA    linear   EST 09-MAR-1998
DEFINITION  zp38d10.r1 Stratagene muscle 937209 Homo sapiens cDNA clone
            IMAGE:611731 5' similar to SW:AAKG_RAT P80385 5'-AMP-ACTIVATED
            PROTEIN KINASE, GAMMA CHAIN ;, mRNA sequence.
ACCESSION   AA178898
VERSION     AA178898.1  GI:1760259
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 413)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
            ,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
            White,Y., Wylie,T., Waterston,R. and Wilson,R.
  TITLE     WashU-NCI human EST Project
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1280   Std Error: 0.00
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 255.
FEATURES             Location/Qualifiers
     source          1. .413
                     /organism="Homo sapiens"
                     /db_xref="GDB:4643570"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:611731"
                     /clone_lib="Stratagene muscle 937209"
                     /tissue_type="muscle"
                     /dev_stage="adult"
                     /lab_host="SOLR (kanamycin resistant)"
                     /note="Organ: skeletal muscle; Vector: pBluescript SK-;
                     Site_1: EcoRI; Site_2: XhoI; Cloned unidirectionally.
                     Primer: Oligo dT. Skeletal muscle from patient with
                     malignant hyperthermia.  Average insert size: 1.0 kb;
                     Uni-ZAP XR Vector; ~5' adaptor sequence: 5' GAATTCGGCACGAG
                     3' ~3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTTT 3'"
BASE COUNT       80 a    109 c    117 g    102 t      5 others
ORIGIN


  Query Match             58.4%;  Score 59;  DB 9;  Length 413;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-06;
  Matches   59;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      43 ctccttcctctaccgcactatccaagatttgggcatcggcacattccgagacttggctg 101
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 CTCCTTCCTCTACCGCACTATCCAAGATTTGGGCATCGGCACATTCCGAGACTTGGCTG 59


RESULT   3
BJ072114
LOCUS       BJ072114                633 bp    mRNA    linear   EST 11-DEC-2001
DEFINITION  BJ072114 NIBB Mochii normalized Xenopus tailbud library Xenopus
            laevis cDNA clone XL096j16 5', mRNA sequence.
ACCESSION   BJ072114
VERSION     BJ072114.1  GI:17502303
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 633)
  AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
            ,Y.
  TITLE     Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .633
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="XL096j16"
                     /clone_lib="NIBB Mochii normalized Xenopus tailbud
                     library"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 25"
BASE COUNT      144 a    162 c    148 g    179 t
ORIGIN


  Query Match             48.3%;  Score 48.8;  DB 10;  Length 633;
  Best Local Similarity   68.0%;  Pred. No. 0.0015;
  Matches   68;  Conservative    0;  Mismatches   32;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           |||||||||||| ||| ||||||||  |   |||||  ||||    ||| |  |    ||
Db     131 CTCAAGTTCCTTCACCTCTTTGGTGACACTCTGCCTAGGCCACAGTTCTTGCAGAAGACC 190

Qy      62 atccaagatttgggcatcggcacattccgagacttggctg 101
           ||||  ||  |||||||||||||| ||| | ||  | || |
Db     191 ATCCTGGAGCTGGGCATCGGCACTTTCAGGGATATAGCAG 230


RESULT   4
BG037921
LOCUS       BG037921                514 bp    mRNA    linear   EST 24-JAN-2001
DEFINITION  dc69g08.y1 NICHD XGC Embl Xenopus laevis cDNA clone IMAGE:3402446
            5' similar to TR:Q9ULX8 Q9ULX8 H91620P. ;, mRNA sequence.
ACCESSION   BG037921
VERSION     BG037921.1  GI:12480506
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 514)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
```

```
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Martha Rebbert, Steven L. Klein, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: Xenopus clones from this library are available
            through the I.M.A.G.E. Consortium/LLNL at: info@image.llnl.gov
            Seq primer: -40RP from Gibco
            High quality sequence stop: 437.
FEATURES             Location/Qualifiers
     source          1. .514
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="IMAGE:3402446"
                     /clone_lib="NICHD XGC Emb1"
                     /tissue_type="embryo (stage 10)"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI;
                     Cloned unidirectionally. Primer: Oligo dT. Average insert
                     size 1.55 kb. Constructed by Life Technologies. Note: This
                     is a Xenopus Gene Collection (XGC) library."
BASE COUNT      159 a    101 c    101 g    153 t
ORIGIN


  Query Match             38.4%;  Score 38.8;  DB 10;  Length 514;
  Best Local Similarity   62.2%;  Pred. No. 0.76;
  Matches   61;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           |||||||||||| ||  | ||||  ||     ||||   |||  |||||  |  | |  | |
Db     283 CTCAAGTTCCTTCAGCTTTTTGTGTCAGAAATGCCAAAGCCTGCCTTCATGAAACAGAAT 342

Qy      62 atccaagatttgggcatcggcacattccgagacttggc 99
             |  | ||  | || ||||||||||| ||   || | ||
Db     343 TTAGAGGAGCTTGGGATCGGCACATACCACAACATTGC 380


RESULT   5
BG027175/c
LOCUS       BG027175                 918 bp    mRNA    linear   EST 24-JAN-2001
DEFINITION  602295858F1 NIH_MGC_86 Homo sapiens cDNA clone IMAGE:4390238 5',
            mRNA sequence.
ACCESSION   BG027175
VERSION     BG027175.1  GI:12415539
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 918)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10079  row: d  column: 15
            High quality sequence stop: 691.
FEATURES             Location/Qualifiers
     source          1. .918
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4390238"
                     /clone_lib="NIH_MGC_86"
                     /tissue_type="osteosarcoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: bone; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.533 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      259 a    138 c    263 g    258 t
ORIGIN


  Query Match             34.5%;  Score 34.8;  DB 10;  Length 918;
  Best Local Similarity   60.6%;  Pred. No. 11;
  Matches   57;  Conservative    0;  Mismatches   37;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           |||  ||| || | | |||||   ||||   | | ||  ||||||    ||| |||| ||
Db     912 CTCCCGTTTCTTCCCCTCTTTTTCTCCCCCTTTCTCCCCCCCTCCCCTTTCTTCCGCCCT 853

Qy      62 atccaagatttgggcatcggcacattccgagact 95
             | ||| ||||   |    | ||| | || | ||
Db     852 TTTCAAAATTTTCCCCCTTGTACACTTCGTGGCT 819


RESULT   6
AW147353/c
LOCUS       AW147353                 451 bp    mRNA    linear   EST 27-JUN-2001
DEFINITION  da01h11.y1 Xenopus laevis oocyte Xenopus laevis cDNA clone
            XENOPUS_SOURCE_ID:xlnoc001d22 5' similar to TR:O70257 O70257
            SYNTAXIN 7. ;, mRNA sequence.
ACCESSION   AW147353
VERSION     AW147353.1  GI:6195249
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 451)
  AUTHORS   Clifton,S., Johnson,S.L., Blumberg,B., Song,J., Hillier,L., Pape,D.
            , Martin,J., Wylie,T., Underwood,K., Theising,B., Bowers,Y., Person
            ,B., Gibbons,M., Harvey,N., Ritter,E., Jackson,Y., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     WashU Xenopus EST project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Other_ESTs: da01h11.x1
            Contact: Sandy Clifton, Ph.D.
            WashU Xenopus EST project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Library constructed by Bruce Blumberg
             Library normalized by Jihwan Song
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: Xenopus clone distribution information for
            this library can be found through Research Genetics, visit their
            web page at: http://www.resgen.com/
            Seq primer: -40RP from Gibco
            High quality sequence stop: 409.
FEATURES             Location/Qualifiers
     source          1. .451
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="XENOPUS_SOURCE_ID:xlnoc001d22"
                     /clone_lib="Xenopus laevis oocyte"
                     /tissue_type="oocyte (stages 5 and 6)"
                     /lab_host="Top-10 F'"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; cDNA was prepared from 2ug of poly A+ RNA.
                     EcoRI-XhoI cut cDNA was then ligated into UniZap-XR
```

(Stratagene) with EcoRI at the 5' end and XhoI at the 3' end. SS-library phagemids were prepared by mass excision from the original library and normalized by hybridization to biotinylated driver (prepared from the same library by PCR) to Cot-omega of 11. After removal of hybrids and excess driver by streptavidin sepharose chromatography, the ss-phagemids were made double stranded and electroporated into Top-10 F'. Original library contruction by Bruce Blumberg (Blumberg et al., 1991 Science 253, 194-196; Hawlet et al., 1995, Genes Dev. 9, 2923-2935). Normalized by Jihwan Song (Song, Cho and Blumberg, unpublished). Note: This is a Xenopus Gene Collection (XGC) library."
BASE COUNT 149 a 95 c 109 g 98 t
ORIGIN

Query Match 31.9%; Score 32.2; DB 9; Length 451;
Best Local Similarity 61.2%; Pred. No. 47;
Matches 52; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 9 tcctgcacatctttggttccctgctgccccggccctccttcctctaccgcactatccaag 68
|| || ||||| |||| || ||| | | ||| |||| ||| | || || |
Db 207 TCTGTCAAATCTTGGGTTTGCTTCTGGCTCTGCCACGCTTCGTCTCCACCATTAAAAATC 148

Qy 69 atttgggcatcggcacattccgaga 93
||| |||| ||| || ||||
Db 147 TTTTCATCATCTGCAAGGTCTGAGA 123

RESULT 7
BJ036397/c
LOCUS BJ036397 470 bp mRNA linear EST 06-DEC-2001
DEFINITION BJ036397 NIBB Mochii normalized Xenopus neurula library Xenopus laevis cDNA clone XL036d24 5', mRNA sequence.
ACCESSION BJ036397
VERSION BJ036397.1 GI:17396982
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae; Xenopodinae; Xenopus.
REFERENCE 1 (bases 1 to 470)
AUTHORS Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara,Y.
TITLE Expressed genes in X. laevis embryo
JOURNAL Unpublished (2001)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES Location/Qualifiers
source 1. .470
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone="XL036d24"
/clone_lib="NIBB Mochii normalized Xenopus neurula library"
/tissue_type="whole embryo"
/dev_stage="stage 15"
BASE COUNT 148 a 108 c 112 g 101 t 1 others
ORIGIN

Query Match 31.9%; Score 32.2; DB 10; Length 470;
Best Local Similarity 61.2%; Pred. No. 47;
Matches 52; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 9 tcctgcacatctttggttccctgctgccccggccctccttcctctaccgcactatccaag 68
|| || ||||| |||| || ||| | | ||| |||| ||| | || || |
Db 296 TCTGTCAAATCTTGGGTTTGCTTCTGGCTCTGCCACGCTTCGTCTCCACCATTAAAAATC 237

Qy 69 atttgggcatcggcacattccgaga 93
||| |||| ||| || ||||
Db 236 TTTTCATCATCTGCAAGGTCTGAGA 212

RESULT 8
AV914030/c
LOCUS AV914030 580 bp mRNA linear EST 18-JAN-2002
DEFINITION AV914030 K. Sato unpublished cDNA library, cv. Haruna Nijo germination shoots Hordeum vulgare subsp. vulgare cDNA clone bags4d24 5', mRNA sequence.
ACCESSION AV914030
VERSION AV914030.1 GI:18209807
KEYWORDS EST.
SOURCE Hordeum vulgare subsp. vulgare.
ORGANISM Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Hordeum.
REFERENCE 1 (bases 1 to 580)
AUTHORS Sato,K., Saisho,D. and Takeda,K.
TITLE Barley EST sequencing project in NIG and Okayama Univ
JOURNAL Unpublished (2002)
COMMENT Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
FEATURES Location/Qualifiers
source 1. .580
/organism="Hordeum vulgare subsp. vulgare"
/cultivar="Haruna Nijo"
/db_xref="taxon:112509"
/clone="bags4d24"
/clone_lib="K. Sato unpublished cDNA library, cv. Haruna Nijo germination shoots"
/tissue_type="shoots"
/dev_stage="germination"
BASE COUNT 165 a 137 c 141 g 137 t
ORIGIN

Query Match 31.9%; Score 32.2; DB 9; Length 580;
Best Local Similarity 61.2%; Pred. No. 50;
Matches 52; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

Qy 9 tcctgcacatctttggttccctgctgccccggccctccttcctctaccgcactatccaag 68
|| || ||||| |||| || ||| | | ||| |||| ||| | || || |
Db 257 TCTGTCAAATCTTGGGTTTGCTTCTGGCTCTGCCACGCTTCGTCTCCACCATTAAAAATC 198

Qy 69 atttgggcatcggcacattccgaga 93
||| |||| ||| || ||||
Db 197 TTTTCATCATCTGCAAGGTCTGAGA 173

RESULT 9
BG101440
LOCUS BG101440 684 bp mRNA linear EST 29-JAN-2001
DEFINITION uy72c12.y1 McCarrey Eddy round spermatid Mus musculus cDNA clone IMAGE:3664919 5' similar to TR:Q9ULX8 Q9ULX8 H91620P. ;, mRNA sequence.
ACCESSION BG101440
VERSION BG101440.1 GI:12596757
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```
REFERENCE   1  (bases 1 to 684)
  AUTHORS   Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
            Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person
            ,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter
            ,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
            Waterston,R. and Wilson,R.
  TITLE     The WashU-NCI Mouse EST Project 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Marra M/WashU-NCI Mouse EST Project 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:1425687
            Seq primer: Primer name ambiguous
            High quality sequence stop: 387.
FEATURES             Location/Qualifiers
     source          1. .684
                     /organism="Mus musculus"
                     /strain="CD-1"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:3664919"
                     /clone_lib="McCarrey Eddy round spermatid"
                     /sex="male"
                     /tissue_type="round spermatids, pooled from multiple mice"
                     /dev_stage="60 day"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: testis; Vector: pBluescript SK+ (Stratagene
                     ); Site_1: XhoII; Site_2: EcoRI; cDNA oligo dT-primed
                     [5'-(GA)10-ACTAGTCTCGAGTTTTTTTTTTTT-3'] and directionally
                     cloned  using 5' linkers 5'-AATTCGGCACGAG-3' and
                     5'-CTCGTGCCG-3'. Size selection of >400bp material gives
                     average insert size ranging from 1-2 kb. Library was mass
                     excised (from  lambda-UniZAP-XR) and resulting
                     single-stranded phagemids were prepped and tranformed
                     into DH10B. Library contains 98.5% recombinants.
                     References: J. Androl. 20:635-639 and Gene 25:263-269.
                     Library constructed and donated by J. McCarrey, Ph.D.
                     (Southwest Foundation for Biomedical Research, Dept. of
                     Genetics); excision done by E.M. Eddy, Ph.D. (National
                     Institutes of  Health, National Institute of Environmental
                     Health Sciences).  Original lambda-based library is
                     available through ATCC, catalog #63423."
BASE COUNT      208 a    163 c    134 g    177 t      2 others
ORIGIN


  Query Match             31.1%;  Score 31.4;  DB 10;  Length 684;
  Best Local Similarity   57.1%;  Pred. No. 87;
  Matches   56;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           |||||||||||  ||  | |||   ||     ||||  ||||  |||| |  | |  |
Db     541 CTCAAGTTCCTCCAGCTTTTTATGTCTGACATGCCAAAGCCCGGCTTCATGAAGCAGAAC 600

Qy      62 atccaagatttgggcatcggcacattccgagacttggc 99
             |  | ||  | || |||||  || | ||   |  | ||
Db     601 CTGTATGAGCTNGGAATCGGAACGTACCACCATATTGC 638


RESULT  10
BM488662
LOCUS       BM488662                  536 bp    mRNA    linear   EST 07-FEB-2002
DEFINITION  pgm2n.pk008.g21 Normalized Chicken Breast Muscle, Leg Muscle, and
            Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cDNA
            clone pgm2n.pk008.g21 5' similar to gb|AAC52580.1 (U42413)
            5'-AMP-activated protein kinase, gamma-1 subunit [Rattus norvegicus
            ], mRNA sequence.
ACCESSION   BM488662
VERSION     BM488662.1  GI:18609593
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 536)
  AUTHORS   Cogburn,L.A. and Monsonego-Ornan,E.
  TITLE     ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and
            Epiphyseal Growth Plate cDNA library, USDA/IFAFS Animal Genome
            Project
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Larry A. Cogburn
            University of Delaware
            Townsend Hall, Newark, DE 19717, USA
            Tel: 302-831-1335
            Fax: 302-831-2822
            Email: cogburn@udel.edu, www.chickest.udel.edu.
FEATURES             Location/Qualifiers
     source          1. .536
                     /organism="Gallus gallus"
                     /strain="Commercial broiler and  Ottawa Res. Centre
                     Strains 90 & 21"
                     /db_xref="taxon:9031"
                     /clone="pgm2n.pk008.g21"
                     /clone_lib="Normalized Chicken Breast Muscle, Leg Muscle,
                     and Epiphyseal Growth Plate cDNA library (pgm2n)"
                     /sex="Male and Female"
                     /tissue_type="Breast muscle, leg muscle and epiphyseal
                     growth plate"
                     /dev_stage="Breast,leg:Embryo(d19);post-hatch(1d,1,3,5,7,9
                     ,11 weeks);growth plate(1d,7d,14d post-hatch)"
                     /lab_host="E. coli EMDH10B"
                     /note="Vector: pCMVSPORT6; Library made from equivalent
                     pools of total RNA isolated from each tissue (embryonic
                     muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth
                     plate 33.3% of the final RNA pool). Single pass sequencing
                     from 5'-end"
BASE COUNT      117 a    171 c    132 g    116 t
ORIGIN


  Query Match             30.9%;  Score 31.2;  DB 10;  Length 536;
  Best Local Similarity   57.0%;  Pred. No. 92;
  Matches   57;  Conservative    0;  Mismatches   43;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           ||||||||||||  |  |||||  | |   | | ||   |||    ||  |   ||| |||
Db     206 CTCAAGTTCCTCAAACTCTTTATTGCAGAGGTCCCAAAGCCTGAGTTTATGGCCCGGACT 265

Qy      62 atccaagatttgggcatcggcacattccgagacttggctg 101
             |  | ||  ||   ||||| || | | |  || | || |
Db     266 TTGGAGGAGCTGCAGATCGGTACCTACAGCAACATCGCCG 305


RESULT  11
BM487789
LOCUS       BM487789                  595 bp    mRNA    linear   EST 07-FEB-2002
DEFINITION  pgm2n.pk005.j24 Normalized Chicken Breast Muscle, Leg Muscle, and
            Epiphyseal Growth Plate cDNA library (pgm2n) Gallus gallus cDNA
            clone pgm2n.pk005.j24 5' similar to gb|AAC52580.1 (U42413)
            5'-AMP-activated protein kinase, gamma-1 subunit [Rattus norvegicus
            ], mRNA sequence.
ACCESSION   BM487789
VERSION     BM487789.1  GI:18608720
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 595)
  AUTHORS   Cogburn,L.A. and Monsonego-Ornan,E.
```

```
  TITLE     ESTs from Normalized Chicken Breast Muscle, Leg Muscle, and
            Epiphyseal Growth Plate cDNA library, USDA/IFAFS Animal Genome
            Project
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Larry A. Cogburn
            University of Delaware
            Townsend Hall, Newark, DE 19717, USA
            Tel: 302-831-1335
            Fax: 302-831-2822
            Email: cogburn@udel.edu, www.chickest.udel.edu.
FEATURES             Location/Qualifiers
     source          1. .595
                     /organism="Gallus gallus"
                     /strain="Commercial broiler and  Ottawa Res. Centre
                     Strains 90 & 21"
                     /db_xref="taxon:9031"
                     /clone="pgm2n.pk005.j24"
                     /clone_lib="Normalized Chicken Breast Muscle, Leg Muscle,
                     and Epiphyseal Growth Plate cDNA library (pgm2n)"
                     /sex="Male and Female"
                     /tissue_type="Breast muscle, leg muscle and epiphyseal
                     growth plate"
                     /dev_stage="Breast,leg:Embryo(d19);post-hatch(1d,1,3,5,7,9
                     ,11 weeks);growth plate(1d,7d,14d post-hatch)"
                     /lab_host="E. coli EMDH10B"
                     /note="Vector: pCMVSPORT6; Library made from equivalent
                     pools of total RNA isolated from each tissue (embryonic
                     muscle 33.3%; juvenile muscle 33.3%; and epiphyseal growth
                     plate 33.3% of the final RNA pool). Single pass sequencing
                     from 5'-end"
BASE COUNT      126 a    199 c    137 g    133 t
ORIGIN


  Query Match             30.9%;  Score 31.2;  DB 10;  Length 595;
  Best Local Similarity   57.0%;  Pred. No. 95;
  Matches   57;  Conservative    0;  Mismatches   43;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           ||||||||||||  |  |||||  | |   | | ||   |||    ||  |   ||| |||
Db     410 CTCAAGTTCCTCAAACTCTTTATTGCAGAGGTCCCAAAGCCTGAGTTTATGGCCCGGACT 469

Qy      62 atccaagatttgggcatcggcacattccgagacttggctg 101
             |  | ||  ||   ||||| || | | |  || | || |
Db     470 TTGGAGGAGCTGCAGATCGGTACCTACAGCAACATCGCCG 509


RESULT  12
BM440762
LOCUS       BM440762                 647 bp    mRNA    linear   EST 01-FEB-2002
DEFINITION  pgr1n.pk002.19 Normalized Chicken Reproductive Tract cDNA Library
            (pgr1n) Gallus gallus cDNA clone pgr1n.pk002.19 5' similar to
            gi|4506061 ref|NP_002724.1| protein kinase, AMP-activated, gamma 1
            non-catalytic subunit; AMPK gamma 1; Protein kinase, AMP-activated,
            noncatalytic, gemma-1 [Homo sapiens] gi|12737489 ref|XP_006778.2|
            protein kinase, AMP-activated, gamma 1, mRNA sequence.
ACCESSION   BM440762
VERSION     BM440762.1  GI:18471537
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 647)
  AUTHORS   Cogburn,L.A. and Nys,Y.
  TITLE     ESTs from Normalized Chicken Reproductive Tract cDNA library-
            University of Delaware and INRA, Tours-Poultry Unit Project
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Larry A. Cogburn
            University of Delaware
            Townsend Hall, Newark, DE 19717, USA
            Tel: 302-831-1335
            Fax: 302-831-2822
            Email: cogburn@udel.edu, www.chickest.udel.edu.
FEATURES             Location/Qualifiers
     source          1. .647
                     /organism="Gallus gallus"
                     /strain="Commercial broiler and layer"
                     /db_xref="taxon:9031"
                     /clone="pgr1n.pk002.19"
                     /clone_lib="Normalized Chicken Reproductive Tract cDNA
                     Library (pgr1n)"
                     /sex="Male and Female"
                     /tissue_type="Testis, ovary and oviduct"
                     /dev_stage="Various stages;embryonic, post-hatch, immature
                     and sexually-mature"
                     /lab_host="E. coli EMDH10B"
                     /note="Vector: pCMVSPORT6; Library made from three total
                     RNA pools from each tissue (testis 25%, ovary 25%, and
                     oviduct 50% of final RNA pool); Single pass sequencing
                     from 5'-end"
BASE COUNT      137 a    222 c    160 g    128 t
ORIGIN


  Query Match             30.9%;  Score 31.2;  DB 10;  Length 647;
  Best Local Similarity   57.0%;  Pred. No. 97;
  Matches   57;  Conservative    0;  Mismatches   43;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           ||||||||||||  |  |||||  | |   | | ||   |||    ||  |   ||| |||
Db     514 CTCAAGTTCCTCAAACTCTTTATTGCAGAGGTCCCAAAGCCTGAGTTTATGGCCCGGACT 573

Qy      62 atccaagatttgggcatcggcacattccgagacttggctg 101
             |  | ||  ||   ||||| || | | |  || | || |
Db     574 TTGGAGGAGCTGCAGATCGGTACCTACAGCAACATCGCCG 613


RESULT  13
AJ395115
LOCUS       AJ395115                 649 bp    mRNA    linear   EST 25-JAN-2001
DEFINITION  AJ395115 dkfz426 Gallus gallus cDNA clone 21c2r1, mRNA sequence.
ACCESSION   AJ395115
VERSION     AJ395115.1  GI:7125706
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 649)
  AUTHORS   Abdrakhmanov,I., Lodygin,D., Geroth,P., Arakawa,H., Law,A., Plachy
            ,J., Korn,B. and Buerstedde,J.M.
  TITLE     A large database of chicken bursal ESTs as a resource for the
            analysis of vertebrate gene function
  JOURNAL   Genome Res. 10 (12), 2062-2069 (2000)
  MEDLINE   20568495
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinistr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES             Location/Qualifiers
     source          1. .649
                     /organism="Gallus gallus"
                     /strain="CB"
                     /db_xref="taxon:9031"
                     /clone="21c2r1"
                     /clone_lib="dkfz426"
                     /tissue_type="Bursa of Fabricius"
BASE COUNT      130 a    227 c    155 g    137 t
ORIGIN


  Query Match             30.9%;  Score 31.2;  DB 9;  Length 649;
```

```
  Best Local Similarity   57.0%;  Pred. No. 97;
  Matches   57;  Conservative    0;  Mismatches   43;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           ||||||||||| |  ||||| | |   | | ||   |||    ||  |   ||| |||
Db     167 CTCAAGTTCCTCAAACTCTTTATTGCAGAGGTCCCAAAGCCTGAGTTTATGGCCCGGACT 226

Qy      62 atccaagatttgggcatcggcacattccgagacttggctg 101
             | | ||  ||   ||||| || | | |  || | || |
Db     227 TTGGAGGAGCTGCAGATCGGTACCTACAGCAACATCGCCG 266


RESULT  14
AJ396118
LOCUS       AJ396118                758 bp    mRNA    linear   EST 25-JAN-2001
DEFINITION  AJ396118 dkfz426 Gallus gallus cDNA clone 25f16r1, mRNA sequence.
ACCESSION   AJ396118
VERSION     AJ396118.1  GI:7127728
KEYWORDS    EST.
SOURCE      chicken.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 758)
  AUTHORS   Abdrakhmanov,I., Lodygin,D., Geroth,P., Arakawa,H., Law,A., Plachy
            ,J., Korn,B. and Buerstedde,J.M.
  TITLE     A large database of chicken bursal ESTs as a resource for the
            analysis of vertebrate gene function
  JOURNAL   Genome Res. 10 (12), 2062-2069 (2000)
  MEDLINE   20568495
COMMENT     Contact: Buerstedde JM
            Cellular Immunology
            Heinrich-Pette-Institute
            Martinistr. 52, 20251 Hamburg, Germany
            Email: URL: http://genetics.hpi.uni-hamburg.de/dt40est.html.
FEATURES             Location/Qualifiers
     source          1. .758
                     /organism="Gallus gallus"
                     /strain="CB"
                     /db_xref="taxon:9031"
                     /clone="25f16r1"
                     /clone_lib="dkfz426"
                     /tissue_type="Bursa of Fabricius"
BASE COUNT      168 a    238 c    186 g    166 t
ORIGIN


  Query Match             30.9%;  Score 31.2;  DB 9;  Length 758;
  Best Local Similarity   57.0%;  Pred. No. 1e+02;
  Matches   57;  Conservative    0;  Mismatches   43;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           ||||||||||| |  ||||| | |   | | ||   |||    ||  |   ||| |||
Db     388 CTCAAGTTCCTCAAACTCTTTATTGCAGAGGTCCCAAAGCCTGAGTTTATGGCCCGGACT 447

Qy      62 atccaagatttgggcatcggcacattccgagacttggctg 101
             | | ||  ||   ||||| || | | |  || | || |
Db     448 TTGGAGGAGCTGCAGATCGGTACCTACAGCAACATCGCCG 487


RESULT  15
H64260
LOCUS       H64260                  413 bp    mRNA    linear   EST 18-OCT-1995
DEFINITION  yr70c12.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone
            IMAGE:210646 5', mRNA sequence.
ACCESSION   H64260
VERSION     H64260.1  GI:1023000
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 413)
  AUTHORS   Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B.,
            Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins
            ,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore
            ,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T.,
            Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E.,
            Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
  TITLE     Generation and analysis of 280,000 human expressed sequence tags
  JOURNAL   Genome Res. 6 (9), 807-828 (1996)
  MEDLINE   97044478
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 933
            High quality sequence stops: 383
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 933    Std Error: 0.00
            Seq primer: M13RP1
            High quality sequence stop: 383.
FEATURES             Location/Qualifiers
     source          1. .413
                     /organism="Homo sapiens"
                     /db_xref="GDB:3783487"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:210646"
                     /clone_lib="Soares fetal liver spleen 1NFLS"
                     /sex="male"
                     /dev_stage="20 week-post conception fetus"
                     /lab_host="DH10B (ampicillin resistant)"
                     /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                     with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                     1st strand cDNA was primed with a Pac I - oligo(dT) primer
                     [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Pac I and cloned into the Pac I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT      132 a     87 c     74 g    119 t      1 others
ORIGIN


  Query Match             30.5%;  Score 30.8;  DB 10;  Length 413;
  Best Local Similarity   57.1%;  Pred. No. 1.1e+02;
  Matches   56;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           ||||||||||| ||  | |||   ||     ||||   |||  ||||| |  | |  |
Db     186 CTCAAGTTCCTCCAGCTTTTTATGTCTGATATGCCAAAGCCTGCCTTCATGAAGCAGAAC 245

Qy      62 atccaagatttgggcatcggcacattccgagacttggc 99
             |  | ||  | || || || || | ||   || | ||
Db     246 CTGGATGAGCTTGGAATAGGAACGTACCACAACATTGC 283


Search completed: October  3, 2002, 16:15:11
Job time: 16966 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 3, 2002, 16:24:51 ; Search time 180.77 Seconds
(without alignments)
137.241 Million cell updates/sec

Title: US-09-826-581-5_COPY_1000_1100
Perfect score: 101
Sequence: 1 gctcaagttcctgcacatct.....cacattccgagacttggctg 101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	30.8	30.5	1435	2	US-08-878-989-14	Sequence 14, Appl
	2	30.8	30.5	1435	4	US-09-272-796-14	Sequence 14, Appl
	3	29.4	29.1	1260	1	US-07-866-979-3	Sequence 3, Appli
	4	29.4	29.1	1260	2	US-08-466-906B-3	Sequence 3, Appli
	5	29.4	29.1	1260	3	US-08-706-281A-3	Sequence 3, Appli
	6	29.4	29.1	1260	4	US-09-201-746-3	Sequence 3, Appli
	7	29.4	29.1	1260	4	US-09-097-231-3	Sequence 3, Appli
	8	27.8	27.5	1378	1	US-08-759-848-2	Sequence 2, Appli
	9	27.8	27.5	1378	5	PCT-US95-09383-2	Sequence 2, Appli
c	10	27.6	27.3	5020	3	US-08-938-291A-3	Sequence 3, Appli
	11	26.4	26.1	5975	1	US-08-336-257A-3	Sequence 3, Appli
c	12	25.6	25.3	1984	2	US-08-822-028-16	Sequence 16, Appl
c	13	25.6	25.3	1984	4	US-08-479-285-16	Sequence 16, Appl
c	14	25.4	25.1	1509	2	US-08-481-337A-1	Sequence 1, Appli
c	15	25.4	25.1	1509	2	US-08-696-268B-1	Sequence 1, Appli
c	16	25.4	25.1	1509	5	PCT-US95-05467-1	Sequence 1, Appli
c	17	25.4	25.1	1596	5	PCT-US94-11328A-3	Sequence 3, Appli
	18	25.4	25.1	1702	4	US-08-934-254-26	Sequence 26, Appl
c	19	25.4	25.1	1984	4	US-09-382-256-1	Sequence 1, Appli
c	20	25.4	25.1	1984	4	US-09-395-115-1	Sequence 1, Appli
c	21	25.4	25.1	1984	4	US-08-436-265-1	Sequence 1, Appli
	22	25.2	25.0	2264	4	US-09-126-109-9	Sequence 9, Appli
	23	25.2	25.0	5962	6	5386025-5	Patent No. 5386025
	24	25.2	25.0	5975	1	US-08-404-354B-1	Sequence 1, Appli
	25	25.2	25.0	5975	1	US-08-314-083B-1	Sequence 1, Appli
	26	25.2	25.0	5975	1	US-08-435-675B-1	Sequence 1, Appli
	27	25.2	25.0	5975	3	US-08-884-599-1	Sequence 1, Appli
c	28	24.8	24.6	1419	1	US-08-103-739B-1	Sequence 1, Appli
c	29	24.8	24.6	1419	1	US-08-474-404-1	Sequence 1, Appli
c	30	24.8	24.6	1419	2	US-08-485-845-1	Sequence 1, Appli
c	31	24.8	24.6	1419	2	US-08-482-714-1	Sequence 1, Appli
c	32	24.8	24.6	1419	4	US-09-211-416-1	Sequence 1, Appli
c	33	24.8	24.6	1419	4	US-09-059-958-1	Sequence 1, Appli
	34	24.4	24.2	261	2	US-08-592-383-7	Sequence 7, Appli
	35	24.4	24.2	1071	1	US-08-612-986-1	Sequence 1, Appli
	36	24.4	24.2	1071	1	US-08-361-806A-1	Sequence 1, Appli
	37	24.4	24.2	1071	5	PCT-US95-16806A-1	Sequence 1, Appli
	38	24.4	24.2	1094	2	US-08-902-294-1	Sequence 1, Appli
	39	24.4	24.2	1094	3	US-09-178-637-1	Sequence 1, Appli
	40	24.4	24.2	2658	2	US-08-592-383-3	Sequence 3, Appli
	41	24.4	24.2	2928	2	US-08-095-728B-3	Sequence 3, Appli
	42	24.4	24.2	2928	5	PCT-US92-02320A-3	Sequence 3, Appli
	43	24.4	24.2	2940	2	US-08-592-383-1	Sequence 1, Appli
	44	24.4	24.2	2940	6	5171671-1	Patent No. 5171671
	45	24.4	24.2	4002	1	US-08-331-488A-1	Sequence 1, Appli

ALIGNMENTS

```
RESULT 1
US-08-878-989-14
; Sequence 14, Application US/08878989
; Patent No. 5885803
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Guegler, Karl G.
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  DISEASE ASSOCIATED PROTEIN
;    TITLE OF INVENTION:  KINASES
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/878,989
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J J
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0321 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1435 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  PENITUT01
```

```
;      CLONE:  1452972
US-08-878-989-14


  Query Match             30.5%;  Score 30.8;  DB 2;  Length 1435;
  Best Local Similarity   57.1%;  Pred. No. 0.11;
  Matches   56;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           ||||||||||| ||  | |||   ||     ||||   |||  ||||| |  | |  |
Db     617 CTCAAGTTCCTCCAGCTTTTTATGTCTGATATGCCAAAGCCTGCCTTCATGAAGCAGAAC 676

Qy      62 atccaagatttgggcatcggcacattccgagacttggc 99
             |  | ||  | || || || || | ||   || | ||
Db     677 CTGGATGAGCTTGGAATAGGAACGTACCACAACATTGC 714


RESULT   2
US-09-272-796-14
; Sequence 14, Application US/09272796
; Patent No. 6207148
;  GENERAL INFORMATION:
;    APPLICANT:  Bandman, Olga
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Corley, Neil C.
;    APPLICANT:  Guegler, Karl G.
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Goli, Surya K.
;    APPLICANT:  Shah, Purvi
;    TITLE OF INVENTION:  DISEASE ASSOCIATED PROTEIN
;    TITLE OF INVENTION:  KINASES
;    NUMBER OF SEQUENCES:  21
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Drive
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/272,796
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/878,989
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J J
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0321 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-855-0555
;      TELEFAX:  415-845-4166
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1435 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  PENITUT01
;      CLONE:  1452972
US-09-272-796-14


  Query Match             30.5%;  Score 30.8;  DB 4;  Length 1435;
  Best Local Similarity   57.1%;  Pred. No. 0.11;
  Matches   56;  Conservative    0;  Mismatches   42;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           ||||||||||| ||  | |||   ||     ||||   |||  ||||| |  | |  |
Db     617 CTCAAGTTCCTCCAGCTTTTTATGTCTGATATGCCAAAGCCTGCCTTCATGAAGCAGAAC 676

Qy      62 atccaagatttgggcatcggcacattccgagacttggc 99
             |  | ||  | || || || || | ||   || | ||
Db     677 CTGGATGAGCTTGGAATAGGAACGTACCACAACATTGC 714


RESULT   3
US-07-866-979-3
; Sequence 3, Application US/07866979
; Patent No. 5532347
;  GENERAL INFORMATION:
;    APPLICANT:  Cone, Roger D
;    APPLICANT:  Mountjoy, Kathleen G
;    TITLE OF INVENTION:  Melanocyte Stimulating Hormone Receptor
;    TITLE OF INVENTION:  and Uses
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Allegretti & Witcoff, Ltd.
;      STREET:  10 South Wacker Drive, Suite 3000
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/866,979
;      FILING DATE:  19920410
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5532347nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  92,154
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-715-1000
;      TELEFAX:  312-715-1234
;      TELEX:  910-221-5317
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1260 base pairs
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  15..959
;    FEATURE:
;      NAME/KEY:  5'UTR
;      LOCATION:  1..14
;    FEATURE:
;      NAME/KEY:  3'UTR
;      LOCATION:  960..1260
US-07-866-979-3


  Query Match             29.1%;  Score 29.4;  DB 1;  Length 1260;
  Best Local Similarity   58.6%;  Pred. No. 0.33;
  Matches   51;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           ||||    | ||| ||||| ||| ||| || || || | | || ||||||   |  || |
Db     363 CTCATTGACGTGCTCATCTGTGGCTCCATGGTGTCCAGTCTCTGCTTCCTGGGCATCATT 422

Qy      62 atccaagatttgggcatcggcacattc 88
```

```
                |||      ||||  ||  |||
Db    423 GCTATAGACCGCTACATCTCCATCTTC 449


RESULT   4
US-08-466-906B-3
; Sequence 3, Application US/08466906B
; Patent No. 5849871
;  GENERAL INFORMATION:
;    APPLICANT:  Cone, Roger D
;    APPLICANT:  Mountjoy, Kathleen G
;    TITLE OF INVENTION:  Melanocyte Stimulating Hormone Receptor
;    TITLE OF INVENTION:  and Uses
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McDonnell Boehnen Hulbert & Berghoff
;      STREET:  300 South Wacker Drive
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/466,906B
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5849871nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  92,154-H
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-913-0001
;      TELEFAX:  312-913-0002
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1260 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    FEATURE:
;      NAME/KEY:  5'UTR
;      LOCATION:  1..14
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  15..959
;    FEATURE:
;      NAME/KEY:  3'UTR
;      LOCATION:  960..1260
US-08-466-906B-3


  Query Match             29.1%;  Score 29.4;  DB 2;  Length 1260;
  Best Local Similarity   58.6%;  Pred. No. 0.33;
  Matches   51;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           ||||    | ||| ||||| ||| ||| || || || | | || ||||||   |  || |
Db     363 CTCATTGACGTGCTCATCTGTGGCTCCATGGTGTCCAGTCTCTGCTTCCTGGGCATCATT 422

Qy      62 atccaagatttgggcatcggcacattc 88
             |||      ||||  ||  |||
Db     423 GCTATAGACCGCTACATCTCCATCTTC 449


RESULT   5
US-08-706-281A-3
; Sequence 3, Application US/08706281A
; Patent No. 6100048
;  GENERAL INFORMATION:
;    APPLICANT:  Cone, Roger D
;    APPLICANT:  Fan, Wei
;    APPLICANT:  Boston, Bruce A
;    APPLICANT:  Kesterton, Robert A
;    APPLICANT:  Lu, Dongsi
;    APPLICANT:  Chen, Wenbiao
;    TITLE OF INVENTION:  Methods and Reagents for Discovering and
;    TITLE OF INVENTION:  Using Mammalian Melanocortin Receptor Agonists and Antagoni
;    TITLE OF INVENTION:  To Modulate Feeding Behavior in Animals
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McDonnell Boehnen Hulbert & Berghoff
;      STREET:  300 South Wacker Drive
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/706,281A
;      FILING DATE:  04-SEP-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6100048nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  96,886
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-913-0001
;      TELEFAX:  312-913-0002
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1260 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    FEATURE:
;      NAME/KEY:  5'UTR
;      LOCATION:  1..14
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  15..959
;    FEATURE:
;      NAME/KEY:  3'UTR
;      LOCATION:  960..1260
US-08-706-281A-3


  Query Match             29.1%;  Score 29.4;  DB 3;  Length 1260;
  Best Local Similarity   58.6%;  Pred. No. 0.33;
  Matches   51;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           ||||    | ||| ||||| ||| ||| || || || | | || ||||||   |  || |
Db     363 CTCATTGACGTGCTCATCTGTGGCTCCATGGTGTCCAGTCTCTGCTTCCTGGGCATCATT 422

Qy      62 atccaagatttgggcatcggcacattc 88
             |||      ||||  ||  |||
Db     423 GCTATAGACCGCTACATCTCCATCTTC 449


RESULT   6
US-09-201-746-3
; Sequence 3, Application US/09201746
; Patent No. 6268221
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Cone, Roger D
;    APPLICANT:  Mountjoy, Kathleen G
;    TITLE OF INVENTION:  Melanocyte Stimulating Hormone Receptor
;    TITLE OF INVENTION:  and Uses
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McDonnell Boehnen Hulbert & Berghoff
;      STREET:  300 South Wacker Drive
;      CITY:  Chicago
;      STATE:  IL
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/201,746
;      FILING DATE:  01-DEC-1998
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6268221nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  92,154-J
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-913-0001
;      TELEFAX:  312-913-0002
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1260 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA to mRNA
;    FEATURE:
;      NAME/KEY:  5'UTR
;      LOCATION:  1..14
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  15..959
;    FEATURE:
;      NAME/KEY:  3'UTR
;      LOCATION:  960..1260
US-09-201-746-3


  Query Match             29.1%;  Score 29.4;  DB 4;  Length 1260;
  Best Local Similarity   58.6%;  Pred. No. 0.33;
  Matches   51;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           ||||    | ||| ||||| ||| ||| || || || | | || ||||||   |  || |
Db     363 CTCATTGACGTGCTCATCTGTGGCTCCATGGTGTCCAGTCTCTGCTTCCTGGGCATCATT 422

Qy      62 atccaagatttgggcatcggcacattc 88
               |||      ||||  ||  |||
Db     423 GCTATAGACCGCTACATCTCCATCTTC 449


RESULT   7
US-09-097-231-3
; Sequence 3, Application US/09097231
; Patent No. 6278038
;  GENERAL INFORMATION:
;       APPLICANT: Cone, Roger D
;                  Chen, Wenbiao
;                  Low, Malcolm J
;       TITLE OF INVENTION: Mammalian Melanocortin Receptor and Uses
;       NUMBER OF SEQUENCES: 22
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
;            STREET: 300 South Wacker Drive
;            CITY: Chicago
;            STATE: Illinois
;            COUNTRY: USA
;            ZIP: 60606
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.25
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/097,231
;            FILING DATE: 12-Jun-1998
;            CLASSIFICATION: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: No. 6278038nan, Kevin E
;            REGISTRATION NUMBER: 35,303
;            REFERENCE/DOCKET NUMBER: 96,886-C
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 312-913-0001
;            TELEFAX: 312-913-0002
;            TELEX: <Unknown>
;  INFORMATION FOR SEQ ID NO: 3:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 1260 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       MOLECULE TYPE: cDNA to mRNA
;       FEATURE:
;            NAME/KEY:  5'UTR
;            LOCATION:  1..14
;       FEATURE:
;            NAME/KEY:  CDS
;            LOCATION:  15..959
;       FEATURE:
;            NAME/KEY:  3'UTR
;            LOCATION:  960..1260
;       SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-097-231-3


  Query Match             29.1%;  Score 29.4;  DB 4;  Length 1260;
  Best Local Similarity   58.6%;  Pred. No. 0.33;
  Matches   51;  Conservative    0;  Mismatches   36;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           ||||    | ||| ||||| ||| ||| || || || | | || ||||||   |  || |
Db     363 CTCATTGACGTGCTCATCTGTGGCTCCATGGTGTCCAGTCTCTGCTTCCTGGGCATCATT 422

Qy      62 atccaagatttgggcatcggcacattc 88
               |||      ||||  ||  |||
Db     423 GCTATAGACCGCTACATCTCCATCTTC 449


RESULT   8
US-08-759-848-2
; Sequence 2, Application US/08759848
; Patent No. 5750826
;  GENERAL INFORMATION:
;    APPLICANT:  Borkowski, Joseph A.
;    APPLICANT:  Strader, Catherine D.
;    APPLICANT:  Hess, John W.
;    APPLICANT:  Chen, Howard Y.
;    APPLICANT:  Trumbauer, Myrna E.
;    TITLE OF INVENTION:  BRADYKININ B2 RECEPTOR MODIFIED
;    TITLE OF INVENTION:  NON-HUMAN ANIMALS
;    NUMBER OF SEQUENCES:  12
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  John W. Wallen III
;      STREET:  126 E. Lincoln Avenue
;      CITY:  Rahway
;      STATE:  New Jersey
```

```
;       COUNTRY:  USA
;       ZIP:  07065-0900
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/759,848
;       FILING DATE:
;       CLASSIFICATION:  800
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/281,393
;       FILING DATE:  27-JUL-1994
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Wallen III, John W.
;       REGISTRATION NUMBER:  35,403
;       REFERENCE/DOCKET NUMBER:  19234
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (908) 594-3905
;       TELEFAX:  (908) 594-4720
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1378 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-08-759-848-2


  Query Match             27.5%;  Score 27.8;  DB 1;  Length 1378;
  Best Local Similarity   65.1%;  Pred. No. 1.2;
  Matches   41;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       4 caagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcactat 63
           || | ||||| |  || | || | ||||||||||| |     | ||  || | |||| ||
Db     816 CATGCTCCTGAATGTCGTGGGCTTCCTGCTGCCCCTGAGTGTCATCACCTTCTGCACGAT 875

Qy      64 cca 66
             ||
Db     876 GCA 878


RESULT  9
PCT-US95-09383-2
; Sequence 2, Application PC/TUS9509383
;  GENERAL INFORMATION:
;     APPLICANT:  Borkowski, Joseph A.
;     APPLICANT:  Strader, Catherine D.
;     APPLICANT:  Hess, John W.
;     APPLICANT:  Chen, Howard Y.
;     APPLICANT:  Trumbauer, Myrna E.
;     TITLE OF INVENTION:  BRADYKININ B2 RECEPTOR MODIFIED
;     TITLE OF INVENTION:  NON-HUMAN ANIMALS
;     NUMBER OF SEQUENCES:  12
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  John W. Wallen III
;       STREET:  126 E. Lincoln Avenue
;       CITY:  Rahway
;       STATE:  New Jersey
;       COUNTRY:  USA
;       ZIP:  07065-0900
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  PCT/US95/09383
;       FILING DATE:
;       CLASSIFICATION:
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  US 08/281,393
;       FILING DATE:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Wallen III, John W.
;       REGISTRATION NUMBER:  35,403
;       REFERENCE/DOCKET NUMBER:  19234
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (908) 594-3905
;       TELEFAX:  (908) 594-4720
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1378 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
PCT-US95-09383-2


  Query Match             27.5%;  Score 27.8;  DB 5;  Length 1378;
  Best Local Similarity   65.1%;  Pred. No. 1.2;
  Matches   41;  Conservative    0;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       4 caagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcactat 63
           || | ||||| |  || | || | ||||||||||| |     | ||  || | |||| ||
Db     816 CATGCTCCTGAATGTCGTGGGCTTCCTGCTGCCCCTGAGTGTCATCACCTTCTGCACGAT 875

Qy      64 cca 66
             ||
Db     876 GCA 878


RESULT  10
US-08-938-291A-3/c
; Sequence 3, Application US/08938291A
; Patent No. 6117673
;  GENERAL INFORMATION:
;     APPLICANT:  Lev, Sima
;     APPLICANT:  Plowman, Gregory D.
;     APPLICANT:  Schlessinger, Joseph
;     TITLE OF INVENTION:  RDGB PROTEINS AND RELATED
;     TITLE OF INVENTION:  PRODUCTS AND METHODS
;     NUMBER OF SEQUENCES:  11
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Lyon & Lyon
;       STREET:  633 West Fifth Street
;       STREET:  Suite 4700
;       CITY:  Los Angeles
;       STATE:  California
;       COUNTRY:  U.S.A.
;       ZIP:  90071-2066
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
;       MEDIUM TYPE:  storage
;       COMPUTER:  IBM Compatible
;       OPERATING SYSTEM:  IBM P.C. DOS 5.0
;       SOFTWARE:  FastSeq
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/938,291A
;       FILING DATE:  September 26, 1997
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  60/027,337
;       FILING DATE:  October 11, 1996
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Warburg, Richard J.
;       REGISTRATION NUMBER:  32,327
;       REFERENCE/DOCKET NUMBER:  228/172
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (213) 489-1600
;       TELEFAX:  (213) 955-0440
;       TELEX:  67-3510
;   INFORMATION FOR SEQ ID NO:  3:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5020 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-938-291A-3


  Query Match             27.3%;  Score 27.6;  DB 3;  Length 5020;
  Best Local Similarity   55.1%;  Pred. No. 2.2;
  Matches   54;  Conservative    0;  Mismatches   44;  Indels    0;  Gaps    0;

Qy       4 caagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcactat 63
           ||  || | |  | | | ||    | || ||     | ||| |||| | |||  | | ||
Db     666 CACCTTGCAGAGCTTGTATGCGCACATGATGGGGAAGACCTGCTTCTTGTACTCCTCGAT 607

Qy      64 ccaagatttgggcatcggcacattccgagacttggctg 101
           |||    | || ||  ||| |   | | | ||||| ||
Db     606 CCAGTTCTCGGACAGGGGCCCCCGCTGGGTCTTGGTTG 569


RESULT  11
US-08-336-257A-3
; Sequence 3, Application US/08336257A
; Patent No. 5726035
;  GENERAL INFORMATION:
;    APPLICANT:  Jay, Scott D
;    APPLICANT:  Ellis, Steven B.
;    APPLICANT:  Harpold, Michael M.
;    APPLICANT:  Campbell, Kevin P.
;    TITLE OF INVENTION:  CALCIUM CHANNEL COMPOSITIONS AND METHODS
;    NUMBER OF SEQUENCES:  8
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brown, Martin, Haller & McClain
;      STREET:  1660 Union Street
;      CITY:  San Diego
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92101-2926
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ Version 1.5
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/336,257A
;      FILING DATE:  07-NOV-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Seidman, Stephanie L.
;      REGISTRATION NUMBER:  33,779
;      REFERENCE/DOCKET NUMBER:  54898
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 238-0999
;      TELEFAX:  (619) 238-0062
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  5975 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  not relevant
;      TOPOLOGY:  not relevant
;    MOLECULE TYPE:  Genomic DNA
;    FEATURE:
;      NAME/KEY:  Coding Sequence
;      LOCATION:  79...5700
;      OTHER INFORMATION:  \product= "Alpha-1 subunit of animal calcium
;      OTHER INFORMATION:  channel"
US-08-336-257A-3


  Query Match             26.1%;  Score 26.4;  DB 1;  Length 5975;
  Best Local Similarity   55.4%;  Pred. No. 6.2;
  Matches   51;  Conservative    0;  Mismatches   41;  Indels    0;  Gaps    0;

Qy       1 gctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcac 60
           |||||| | | | | |  | | |  || |||||||  | || || | || ||  |  ||
Db    1734 GCTCAACTCCATCCGCTCCATCGCCTCGCTGCTGCTGCTGCTCTTCCTCTTCATCATCAT 1793

Qy      61 tatccaagatttgggcatcggcacattccgag 92
             ||       ||||||||      ||| | |
Db    1794 CTTCGCCCTGCTGGGCATCCAGCTCTTCGGGG 1825


RESULT  12
US-08-822-028-16/c
; Sequence 16, Application US/08822028
; Patent No. 5993813
;  GENERAL INFORMATION:
;    APPLICANT:  MEZES, PETER S
;    APPLICANT:  GOURLIE, BRIAN B
;    APPLICANT:  RIXON, MARK W
;    APPLICANT:  ANDERSON, WH KERR
;    APPLICANT:  KAPLAN, DONALD A
;    APPLICANT:  SCHOLOM, JEFFREY
;    TITLE OF INVENTION:  A NOVEL FAMILY OF HIGH AFFINITY,
;    TITLE OF INVENTION:  MODIFIED ANTIBODIES FOR CANCER TREATMENT
;    NUMBER OF SEQUENCES:  74
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  DUANE C ULMER
;      STREET:  P.O. BOX 1967
;      CITY:  MIDLAND
;      STATE:  MICHIGAN
;      COUNTRY:  USA
;      ZIP:  48641-1967
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/822,028
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/040,687
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  ULMER, DUANE C
;      REGISTRATION NUMBER:  34,941
;      REFERENCE/DOCKET NUMBER:  C-37,075C
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (517) 636-8104
;  INFORMATION FOR SEQ ID NO:  16:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1984 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
US-08-822-028-16


  Query Match             25.3%;  Score 25.6;  DB 2;  Length 1984;
  Best Local Similarity   59.7%;  Pred. No. 7.9;
  Matches   43;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           || |  | |||  |||  |||  |||||  || | |    ||  || ||||      | |
Db     897 CTGAGATCCCTAGACAGATTTATTTCCCAACTTCTCTCAGCCGGCTCCCTCAGGGCAAAT 838

Qy      62 atccaagatttg 73
           |||||||||| |
Db     837 ATCCAAGATTAG 826


RESULT  13
US-08-479-285-16/c
```

```
; Sequence 16, Application US/08479285
; Patent No. 6207815
; GENERAL INFORMATION:
;   APPLICANT: MEZES, PETER S
;   APPLICANT: GOURLIE, BRIAN B
;   APPLICANT: RIXON, MARK W
;   APPLICANT: ANDERSON, WH KERR
;   APPLICANT: KAPLAN, DONALD A
;   APPLICANT: SCHOLOM, JEFFREY
;   TITLE OF INVENTION: A NOVEL FAMILY OF HIGH AFFINITY,
;   TITLE OF INVENTION: MODIFIED ANTIBODIES FOR CANCER TREATMENT
;   NUMBER OF SEQUENCES: 74
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: DUANE C ULMER
;     STREET: P.O. BOX 1967
;     CITY: MIDLAND
;     STATE: MICHIGAN
;     COUNTRY: USA
;     ZIP: 48641-1967
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/479,285
;     FILING DATE: 07-JUN-1995
;     CLASSIFICATION: 536
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/040687
;     FILING DATE: 31-MAR-1993
;   ATTORNEY/AGENT INFORMATION:
;     NAME: ULMER, DUANE C
;     REGISTRATION NUMBER: 34,941
;     REFERENCE/DOCKET NUMBER: C-37,075C
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (517) 636-8104
; INFORMATION FOR SEQ ID NO: 16:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1984 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
US-08-479-285-16


  Query Match             25.3%;  Score 25.6;  DB 4;  Length 1984;
  Best Local Similarity   59.7%;  Pred. No. 7.9;
  Matches   43;  Conservative    0;  Mismatches   29;  Indels    0;  Gaps    0;

Qy       2 ctcaagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcact 61
           || |  | |||  |||  |||  |||||  || | |    ||  || ||||      | |
Db     897 CTGAGATCCCTAGACAGATTTATTTCCCAACTTCTCTCAGCCGGCTCCCTCAGGGCAAAT 838

Qy      62 atccaagatttg 73
           ||||||||||  |
Db     837 ATCCAAGATTAG 826


RESULT  14
US-08-481-337A-1/c
; Sequence 1, Application US/08481337A
; Patent No. 5863738
; GENERAL INFORMATION:
;   APPLICANT: TEN DIJKE, Peter
;   APPLICANT: HELDIN, Carl-Henrik
;   APPLICANT: MIYAZONO, Kohei
;   APPLICANT: SAMPATH, Kuber T.
;   TITLE OF INVENTION: Morphogenic Protein-Specific Cell
;   TITLE OF INVENTION: Surface Receptors and Uses Therefor
;   NUMBER OF SEQUENCES: 18
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Testa, Hurwitz & Thibeault
;     STREET: 125 High St.
;     CITY: Boston
;     STATE: MA
;     COUNTRY: USA
;     ZIP: 02110
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/481,337A
;     FILING DATE: 02-JUN-1995
;     CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;     NAME: MEYERS, Thomas C.
;     REGISTRATION NUMBER: 36,989
;     REFERENCE/DOCKET NUMBER: CRP-097CP2
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (617) 248-7000
;     TELEFAX: (617) 248-7100
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1509 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;     NAME/KEY: CDS
;     LOCATION: 1..1509
;     OTHER INFORMATION: /product= "Human ALK1"
US-08-481-337A-1


  Query Match             25.1%;  Score 25.4;  DB 2;  Length 1509;
  Best Local Similarity   64.4%;  Pred. No. 8.5;
  Matches   38;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       4 caagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcacta 62
           |||||||| |||   |   |||||  |  |  ||  ||||||| || |  | ||||||
Db     216 CAAGTTCCCGCAGCCCCGATGTTCCTGGGGGTGCCTCCCCTCCTCCCGCACCAGCACTA 158


RESULT  15
US-08-696-268B-1/c
; Sequence 1, Application US/08696268B
; Patent No. 5968752
; GENERAL INFORMATION:
;   APPLICANT: ICHIJO, HIDENORI
;   APPLICANT: NISHITOH, HIDEKI
;   APPLICANT: SAMPATH, KUBER T.
;   TITLE OF INVENTION: NOVEL SIGNALLING RECEPTOR FOR
;   TITLE OF INVENTION: MORPHOGENIC PROTEINS
;   NUMBER OF SEQUENCES: 8
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Testa, Hurwitz & Thibeault
;     STREET: 125 High St.
;     CITY: Boston
;     STATE: MA
;     COUNTRY: USA
;     ZIP: 02110
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/696,268B
;     FILING DATE:
;     CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;     NAME: MEYERS, Thomas C.
```

```
;       REGISTRATION NUMBER:  36,989
;       REFERENCE/DOCKET NUMBER:  CRP-117
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (617) 248-7000
;       TELEFAX:  (617) 248-7100
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  1509 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  1..1509
;       OTHER INFORMATION:  /product= "Human ALK-1"
US-08-696-268B-1


  Query Match             25.1%;  Score 25.4;  DB 2;  Length 1509;
  Best Local Similarity   64.4%;  Pred. No. 8.5;
  Matches   38;  Conservative    0;  Mismatches   21;  Indels    0;  Gaps    0;

Qy       4 caagttcctgcacatctttggttccctgctgccccggccctccttcctctaccgcacta 62
           |||||||| |||   |    |||||  |  |  ||  ||||||| || |  | ||||||
Db     216 CAAGTTCCCGCAGCCCCGATGTTCCTGGGGGTGCCTCCCCTCCTCCCGCACCAGCACTA 158


Search completed: October  3, 2002, 16:24:54
Job time: 17134 sec
```